UNIVERSITY OF CALIFORNIA
RIVERSIDE

Agent Based Modeling of Land Use Change: The Case of Shade Coffee in Mexico.

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Environmental Sciences

by

Raymundo Marcos Martinez

March 2014

Dissertation Committee:
Dr. Kenneth A. Baerenklau, Chairperson
Dr. Kurt Schwabe
Dr. Keith C. Knapp

The Dissertation of Raymundo Marcos Martinez is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMETS

I would like to express my sincerest gratitude and appreciation to the following persons:

Dr. Kenneth A. Baerenklau	Without your advice and help I would have never made it through graduate school.
Dr. Kurt Schwabe	You are one of the best professors I have ever met and you were always there to help me when I needed.
Dr. Keith Knapp	Through your example and research you contributed significantly to my education.
Dr. Edward A. Ellis	You help me to start my research career in land use analysis.

Also to UC-MEXUS and CONACYT for all the support and guidance received from the wonderful staff working at those institutions.

To my family in Mexico, and in Thailand, you have made my life happier and meaningful, especially you my lovely Kaew.

ABSTRACT OF THE DISSERTATION

Agent Based Modeling of Land Use Change: The Case of Shade Coffee in Mexico.

by

Raymundo Marcos Martinez

Doctor of Philosophy, Graduate Program in Environmental Sciences
University of California, Riverside, March 2014
Professor Kenneth A. Baerenklau, Chairperson

This research focuses on addressing methodological issues that impact the performance of spatially explicit discrete choice agent-based land use models that are estimated with remotely sensed data. The empirical setting considers land use transitions between agroforests, perennial crops, grass and corn, and fallow lands during the period 1984 – 2006 in a Mexican coffee growing region in which relatively high deforestation rates were observed. As a starting point, a Mixed Conditional – Multinomial Logit model is implemented to highlight assumptions and limitations associated with this standard modeling approach. The results indicate that this model produces theoretically inconsistent parameter estimates for the revenue variable associated with three out of four land uses considered in the analysis. To investigate whether those counterintuitive marginal effects are generated from misclassified land use data, a Latent Multinomial Logit (LMNL) model is implemented. This approach allows the identification of land

use observations that have a high likelihood of being wrongly classified. A reconfiguration of the dataset based on the LMNL model increased the magnitudes of the marginal effects of the analyzed land use drivers in the theoretically expected directions. Next, because static land use models require limiting assumptions that potentially oversimplify the behavioral process followed by landowners, a structural dynamic discrete choice model of land use decisions is implemented under the assumption that land managers are forward-looking and act to maximize their discounted flow of current and future expected utility within a stochastic environment. A comparison between static and dynamic models shows that the directions of the marginal effects corresponding to time-invariant parcel-specific variables generally have the expected directions independent of the selected modeling approach. More importantly, the marginal effect estimates for the revenue variables of the agroforestry and perennial crops categories have the expected direction in the dynamic model. By contrast the myopic modeling approaches generate counter-intuitive results for the revenue variable that corresponds to perennial crops production, which affects the validity of those results for policy design. Finally, a policy simulation exercise shows the sensitivity of welfare estimates to the discount factor selected as representative of the true value used by decision makers.

Table of contents

List of tables

List of figures

Chapter 1
Introduction

The purpose of this brief introductory chapter is to provide a general justification of the relevance of the study of agent based land use decisions, as well as a description of the organization of this thesis. Literature reviews and statements of the research problems that support and motivate the modeling approaches implemented in this research are left to subsequent chapters.

During the last few decades, human-driven alterations of the surface of the Earth have affected relevant patterns and process of the global ecosystem and constitute one of the main determinants of global environmental change (Bonan, DeFries, Coe, & Ojima, 2004; Lambin et al., 2001). It is estimated that in the postindustrial era, land use and land cover change (LULCC) such as conversion of natural to agricultural ecosystems, drainage of wetlands, or biomass burning have generated carbon emissions equivalent to 50% of the global fossil fuel combustion related discharges (Lal, 2004). On the other hand, greenhouse gas emissions from LULCC dynamics exceed that generated by the transportation sector worldwide (Myers Madeira, 2008). Land use management can also modify surface characteristics and affect climate conditions. For instance, overgrazing and agricultural intensification may reduce vegetation and evapotranspiration generating warmer surfaces in areas with low soil moisture content (Bofnan et al. 2004) as well as soil exposure and soil compaction which may result in lower infiltration rates, increased runoff, and drier soils (USDA, 2008).

Landscape dynamics also affect water resources (see Guo, Ma, Yang, & He, 2010; Peierls, Caraco, Pace, & Cole, 1991 for some examples), biodiversity (Haines-Young, 2009; Hansen, DeFries, & Turner, 2004; Maestas, Knight, & Gilgert, 2003) and might alter the frequency or severity of natural disasters such as wildfires (Millington et al., 2008), landslides (Begueria, 2006), floods (Jiang et al., 2008), and hurricanes (Zia, 2012). During the coming decades, significant increases are expected in global population, urban areas, and living standards in developing countries. Those situations will create pressure to further convert or intensify the management of existing ecosystems to satisfy human needs. Given the relevance that the anthropogenic landscape configuration has over key aspects of the global ecosystem, several techniques and methodological approaches have been implemented during the last few decades to identify, analyze and simulate landscape dynamics (see Agarwal, Green, Grove, Evans, & Schweik, 2002; Bürgi, Hersperger, & Schneeberger, 2004; Matthews, Gilbert, Roach, Polhill, & Gotts, 2007; Parker, Manson, Janssen, Hoffmann, & Deadman, 2003 for a review of modelling approaches). The implemented LULCC modeling efforts have facilitated the identification of underlying drivers of landscape dynamics (Bürgi et al., 2004; Geist & Lambin, 2002; Redman, Grove, & Kuby, 2004), and the assessment of potential impacts of human-environment interactions on the global ecosystem .

Nevertheless, there are still plenty of methodological issues in the agent-based land use change literature that require further research. Some assumptions or methodological approaches that directly impact the direction, magnitude and/or statistical significance of parameter estimates do not appear to have enough empirical justification.

For instance, forward-looking behavior, subjective expectations, consumption constraints, neighborhood effects on technology adoption, and risk preferences have received relatively little attention in the land use and land cover change literature.

Given the relevant contribution of environmental services provided by forested areas, a significant amount of research has been done to understand how driving forces of deforestation reconfigure pristine landscapes (Andersen, 1996; Chomitz & Gray, 1996; Geist & Lambin, 2002; Puri, 2006). Unfortunately, the continued pressure over this type of natural resource has reduced the available area of forests stands. The growing recognition that agroforestry production systems can provide forest-like services has motivated a surge of studies of landscape dynamic in agroforestry regions (see Kursten 2000; Blackman & Ávalos-Sartorio 2010; Bhagwat et al. 2008; Shanker & Solanki 2000; Dinata Putra et al. 2005; Swallow et al. 2006; Huang et al. 2002; Schroth 2004 for some examples). Given the socioeconomic and ecological relevance that coffee production has in many developing nations, the empirical applications of the modeling approaches tested in this thesis are based on spatially explicit information collected in a Mexican coffee growing area that has experienced high rates of tree canopy loss during recent years.

There are practical and methodological justifications to extend the analysis of landscape dynamics in agroforestry production areas. From a practical standpoint, an improved understanding of the patterns and processes that generate LULCC in those regions and the associated socioeconomic and environmental consequences would be beneficial for policy design that could help to increase agricultural productivity, provide support in biodiversity conservation efforts, help to control soil degradation and water pollution, facilitate sustainable urbanization processes, and even help to reduce rural poverty and migration in agroforestry areas. From a methodological standpoint the implementation of modeling approaches to incorporate agent's subjective revenue expectation processes; to reduce the impacts of misclassified, incomplete or latent land use information; and to estimate a structural model of land use decisions with forward looking agents[1] and potentially reversible land use decisions, ultimately could lead to better understanding of landscape configuration processes and their drivers. Both issues are considered in this research by using a latent multinomial logit model to identify misclassified land use observations, and by implementing a structural dynamic model of land use choices under uncertainty.

To test alternative modeling options to handle the aforementioned issues, I use remotely sensed data, socioeconomic information and parcel specific characteristics to study land use decisions in a Mexican coffee-growing region during the period 1984 - 2006. During that time window, a substantial amount of tree canopy was lost as farmers moved out of shade coffee and into other crops like corn, citrus and banana. A detailed description of the study site, and the dataset used in the land use models is presented in chapter 2. To highlight the methodological limitations of standard static discrete choice models, chapter 3 describes the implementation of a mixed Multinomial – Conditional logit model to estimate marginal effects of spatial and economic factors that affect land use decisions in the study area. Chapter 4 presents an application of the Latent Multinomial Logit (LMNL) methodology (Caudill, 2006), which is a post-classification

[1] In this thesis agents, decision makers, landowners and land managers are considered equivalent concepts.

procedure that can be used to identify misclassified land use data. The results indicate that the reclassification of the parcels based on the LMNL model increases the magnitudes of the marginal effects of changes in the modeled land use drivers in the theoretically expected direction. To investigate the structure of the behavioral process followed by landowners in the study region, in the last chapter, I implement an agent-based model that consider that the land use choices are dependent on stochastic future returns expectations and parcel characteristics within a framework that allows for the possibility that land use decisions can be reversed. The approach is based on a dynamic discrete choice model that uses the nested fixed point algorithm to compute the maximum likelihood parameter estimates that better explain the decision making process under analysis.

Chapter 2
Description of the study region and dataset.

2.1. Study region.

Deforestation in Mexico has been a relevant issue during the last decades, particularly because this country has been considered one of the five richest biodiversity hotspots in the world (Mas et al., 2004), and because the agriculture and forest sector despite representing only 3.8% of the Gross Domestic Product, employs around 13% of the Mexican labor force representing around 3.3 million farmers that live mostly in poor rural areas (Conservation International, 2013). According to the Food and Agriculture Organization (FAO, 2010) during the period 1990-2000 Mexico registered one of the ten largest annual net losses of forested areas worldwide. Despite the declining trend in the deforestation rates observed after 2000, the protection of the remaining forested areas and the restoration of degraded lands still is a relevant policy issue. According to the Mexican Ministry of Environmental and Natural Resources, one of the Mexican states in which deforestation has significantly reconfigured the landscape is Veracruz. In that state, it is estimated that around 75% of the forested areas have been cleared or significantly degraded (SEMARNAT, 2005).

Tree canopy removal has not only happened in primary or secondary forests, but also in agroforestry regions. Particularly in shade-grown coffee plantations during the 1990's and first half of the 2000's as a result of the long term sustained decline in coffee prices observed during that period. The elimination of the Mexican Institute of Coffee

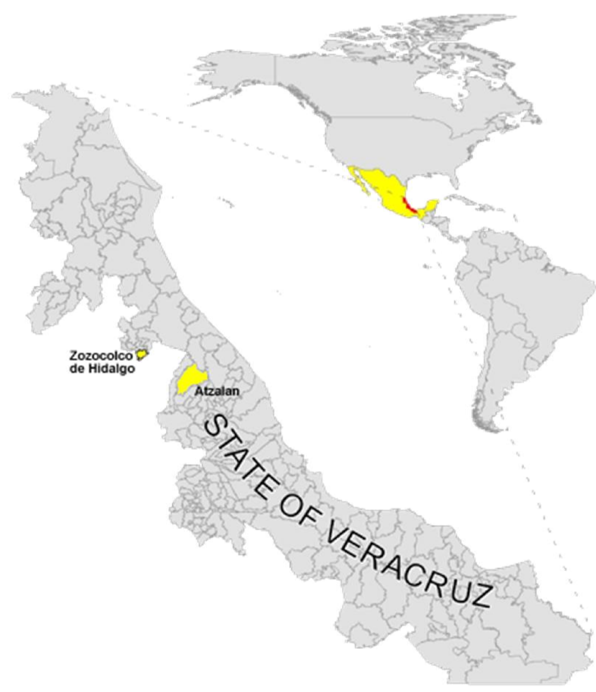
(INMECAFE), which was a governmental agency that subsidized the production, processing, and commercialization of coffee beans, left coffee growers alone in a reconfigured international market (Perfecto, Vandermeer, Mas, & Pinto, 2005). The situation was worst for coffee growers in low-land regions[2] that are not suitable to produce high quality coffee, but that were devoted to this agricultural activity during the period of high coffee prices and heavy governmental subsidies to coffee related activities during the 1970's and 1980's. Landowners in those regions were more prone to experience reconversion of coffee plantations to monocrops (e.g., citrus, banana, cornfields) to grasslands; or to migrate and abandon their lands (Gay, Estrada, Conde, Eakin, & Villers, 2006; Perfecto et al., 2005; Romero, Houston, & Epperson, 2006).

After real coffee prices received by farmers reached their lowest level during the last 100 years in 2002 (Perfecto et al., 2005) the government of the state of Veracruz joined efforts with the Universidad Veracruzana (the main public university in Veracruz), the Common Fund for Commodities, and the International Coffee Organization to implement a pilot project aimed at providing additional income sources and reducing migration rates and tree canopy removal in low land coffee growing areas. The objective was to provide funding and technology transfer to help farmers establish and manage diversified agroforestry production systems. Two regions in the central area of the state of Veracruz were selected to implement the pilot project: Zozocolco, and the low altitude areas of the municipality of Atzalan (see figure 2.1). Those regions have similar agro-ecological characteristics but different socioeconomic and cultural conditions. A

[2] In this study low-lands are defined to be less than 800 meters above sea level.

preliminary version of the dataset constructed as part of this thesis was used in two previous studies of those regions by Ellis et al. (2010) and Baerenklau et al. (2012). Those authors find that the population living in Atzalan is relatively more educated and better connected to commodity markets compared with the population of Zozocolco that has higher poverty rates, is mainly composed by indigenous people, and is relatively isolated from regional commodity markets. Landscape metric and econometric analyses implemented by those authors indicate that coffee-growing areas in Atzalan registered a significant loss of tree canopy, mainly in coffee growing areas in response to the long-term decline in the profitability of that crop. On the other hand, those authors report that coffee growers in Zozocolco either abandoned their plantations or increased the number of allspice trees in their parcels instead of removing the tree canopy to substitute coffee for other crops.

Figure 2.1 Location of Zozocolco de Hidalgo and Atzalan, Veracruz, Mexico.



Since agents' land use decisions in Zozocolco appear to be weakly related to variations in commodity prices and potentially affected by subsistence constraints that require modeling assumptions that differ from the approaches that are implemented in this thesis, I decided to focus my analysis on the study of agent based land use decisions with empirical data corresponding to the low altitude area of Atzalan. Figure 2.2 shows the geographic location of the study region used in this research.

Figure 2.2 Low altitude coffee growing region in Atzalan, Veracruz, Mexico.

The study area is defined by a contiguous geographical surface integrated by around 25,500 hectares distributed across an altitudinal gradient that extends from 80 to 760 meters above sea level. The landscape in that region has gradually reconfigured from secondary forest and coffee parcels to grasslands, citrus groves and banana plantations.

Information collected in 2006 by the Mexican government (SAGARPA, 2006) indicates that at the municipality level citrus production was the main agricultural activity accounting for 68% of the agricultural GDP in the whole municipality. Banana plantations contributed 12% of the production value; corn generated 9% and coffee production after representing the main income source in the region during the previous decades only contributed 5% in that year. On aggregate around 89% of the agricultural GDP in the municipality is generated by agricultural systems that do not require tree canopy, which is an indicator of the landscape degradation experienced in that region. Statistics collected in 2000 indicate that around 82% of the labor force in Atzalan was employed in agricultural activities and that only two thirds of those workers received wages while the remaining arguable worked in their own farms (INEGI, 2001). Additionally, according to CONAPO (2006) Atzalan is considered a municipality with high poverty levels.

2.2. Dataset description.

2.2.1. Land use information

In the development of land use change analysis one of the first challenges is to collect spatially explicit data in a cost-effective way. Depending on the scale of the study, field research (e.g., participatory mapping, surveys), or remotely sensed data can be used to produce land use information to identify landscape dynamics in a particular region. For instance, field research can be used to generate a detailed description of ecological, geographic, and economic landscape-related features at local or regional scales. Nevertheless, its implementation, validation and analysis can be costly when compared with alternatives that rely on remotely sensed data (Cornwall & Jewkes, 1995). Specifically, remote sensing techniques can be used to acquire land use and land cover data of large areas or over inaccessible regions by detecting the natural radiation emitted or reflected by materials at the surface of the earth (Comber et al 2008). There are pros and cons associated with implementing either of these mechanisms, and usually a combination of both approaches results in a more accurate land use/land cover classification. For instance, the information captured by remote sensors may not be enough to distinguish between land features with similar spectral signature (Comber et al. 2008), but groundtruthing verification can help to produce a more accurate landscape analysis. In this thesis remotely sensed data is used with groundtruthing points to produce a time series dataset of land use decisions between relevant land use classes.

One of the implicit assumptions in economic models of land use change is that the period for which the researcher has information corresponds to the period in which agents make their decisions (Rust, 1994). This assumption is rarely recognized but almost general in studies that investigate land use decisions with cross-sectional or time series data, particularly in the analysis based on remotely sensed information. In those types of analyses it is commonly assumed that each of the land use decisions observed at time t was made during that specific period as a function of the explanatory variables observed at that specific point in time. Unfortunately, land use analyses that require remotely sensed information are heavily constrained by the availability of aerial or satellite imagery. In particular, technical and financial constraints limit the collection of historical land use information to only certain points in time. This constitutes a potential issue in the analysis of landscape dynamics since studies based on land use information collected at points in time that are separated by many years are more likely to erroneously model observed land use decisions as being determined in the observed periods, when in fact those choices may have been made at a time for which there is not information in the dataset. Therefore, land use analysis that rely on observations separated by many years may produce inaccurate estimations by modeling relationships between land use decisions and land use drivers that are not temporally linked, especially during periods of relatively large fluctuations in the explanatory variables.

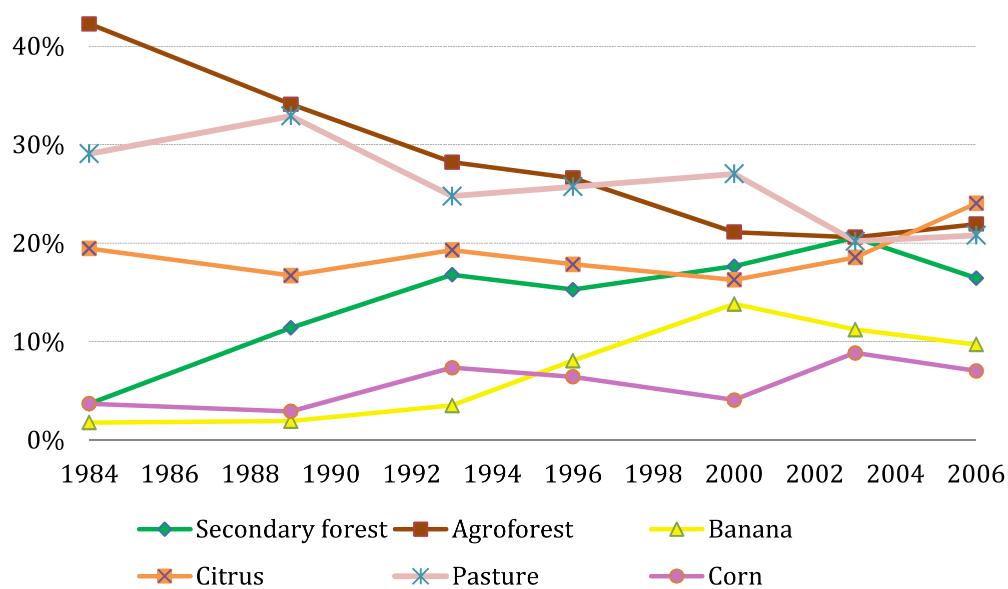
In an effort to reduce the potential estimation problems generated by the aforementioned issue, I collected remotely sensed data for as many periods as possible for the geographical region under analysis. I obtained seven LandSat images for the years 1973, 1984, 1989, 1993, 1996, 2000, 2003, and one Spot image for the year 2006. After masking out cloud-contaminated data present in the satellite images, I used Image Analysis for ArcGIS to implement maximum likelihood classification algorithms to group the pixels in the remotely sensed data into six land use categories: secondary forests, shade grown coffee, banana, citrus, pasture, and corn.

Unfortunately, the spectral information contained in the LandSat image collected in 1973 does not allow separating the pixels into the six mentioned categories. The coarse pixel data in that satellite scene was categorized into forest/agroforestry, grasslands and agricultural areas. Given the differences in the land use classification implemented in the 1973 LandSat image, the land use analysis is limited to the period 1984 – 2006. This will also help to reduce the maximum time separation between observations because several land use transitions are possible during the 11 years between the periods 1973 and 1984.

Figure 2.3 shows how the landscape configuration in the study region changed during the period 1984 – 2006. We can observe a continuous decline in the percentage of land covered by coffee plantations during the period 1984 – 2000 and then the proportion seems to stabilize around 21% during the period 2000 – 2006. Similarly, land devoted to cattle ranching activities presents a declining trend during the period under analysis. Figure 2.3 also allows us to observe that the percentage of land detected as citrus plantations is almost constant during the period of analysis. In contrast, the proportion of banana plantations gradually increased from 2% in 1984 to 13% during 2000. The proportion of cornfields detected in the land use classification do not vary significantly, in 1984 the land devoted to corn production represented 3% of the total surface, that percentage increased to 7% at the end of the observed period. Land detected as secondary forest presents a surprising trend since the relative proportion increased from 4% in 1984 to 18% in 1993, and after reaching the highest percentage for this land use category at 21% in 2003, the proportion ended at 17% during 2006.

Unfortunately, the trend corresponding to secondary forests does not seem to correspond to the landscape configuration observed in the study area. A potential explanation to that trend is that the biomass density commonly found in shade-grown coffee plantations produces spectral signatures that may be difficult to differentiate from pixel information generated by secondary forests. In fact it is highly likely that a significant percentage of the parcels classified as Secondary forest are in fact Agroforest parcels. Unfortunately, the resolution of the LandSat and Spot images do not allow us to accurately discriminate between those two land use classes. We could implement participatory surveys or geo-reference the forested zones in the region to isolate areas that are in fact secondary forest but unfortunately those activities are too costly and time consuming to be implemented in this research.

Figure 2.3. Landscape configuration dynamics in the study region 1984 – 2006



To reduce potential classification errors, and the dimensionality of the modeling approaches that are implemented in this thesis, I integrated the land use classes obtained with the supervised classifiers into three general land use categories: agroforestry (AG) composed of shade grown coffee and secondary forests; grassland and corn for grain (GC), and perennial crops (PC) composed of pixels classified as citrus or banana. Those categories are composed of land uses with similar tree canopy density, profitability and conversion costs. Figures 2.4 and 2.5 show the land use maps generated with the remotely sensed data using the aggregated land use classifications, and figure 2.6 presents the trends observed during the study period. One of the drawbacks of the aggregation of the land use data into the AG, GC and PC classes is that some of the observed trends in figure 2.3 cancel out once we merge the components of each aggregated land use class. For instance, merging areas identified as shade grown coffee plantations or secondary forests into the AG category results in a lower rate of decline than the observed for agroforestry plantations (although as mentioned before those trends are likely affected by misclassified observations).

2.2.2. Spatial autocorrelation and sample selection

The raster images containing the land use classification produced with the remotely sensed data are integrated by around 7,400,000 pixels classified into one of the three broad land use categories mentioned in the previous section. Commonly in the land use literature each pixel is considered an observation of the independent variable, and ideally we should be able to use all the available information to analyze landscape dynamics in the study region. Nevertheless land use decisions are not arbitrary since agents take into account spatially dependent geographic variables as well as neighboring agents' choices to make their decisions. Therefore, we should expect that land use information at the pixel level does not satisfy the assumption that the observations are independent from each other. In fact spatial autocorrelation and spatial heterogeneity are inherent components of land use analysis that need to be controlled to correctly estimate standard errors, and to produce accurate inferences and efficient parameter estimates (Brady & Irwin, 2011; Griffith, 2009; Schnier & Felthoven, 2011).

Figure 2.4. Land use maps low land Atzalan 1973-1993

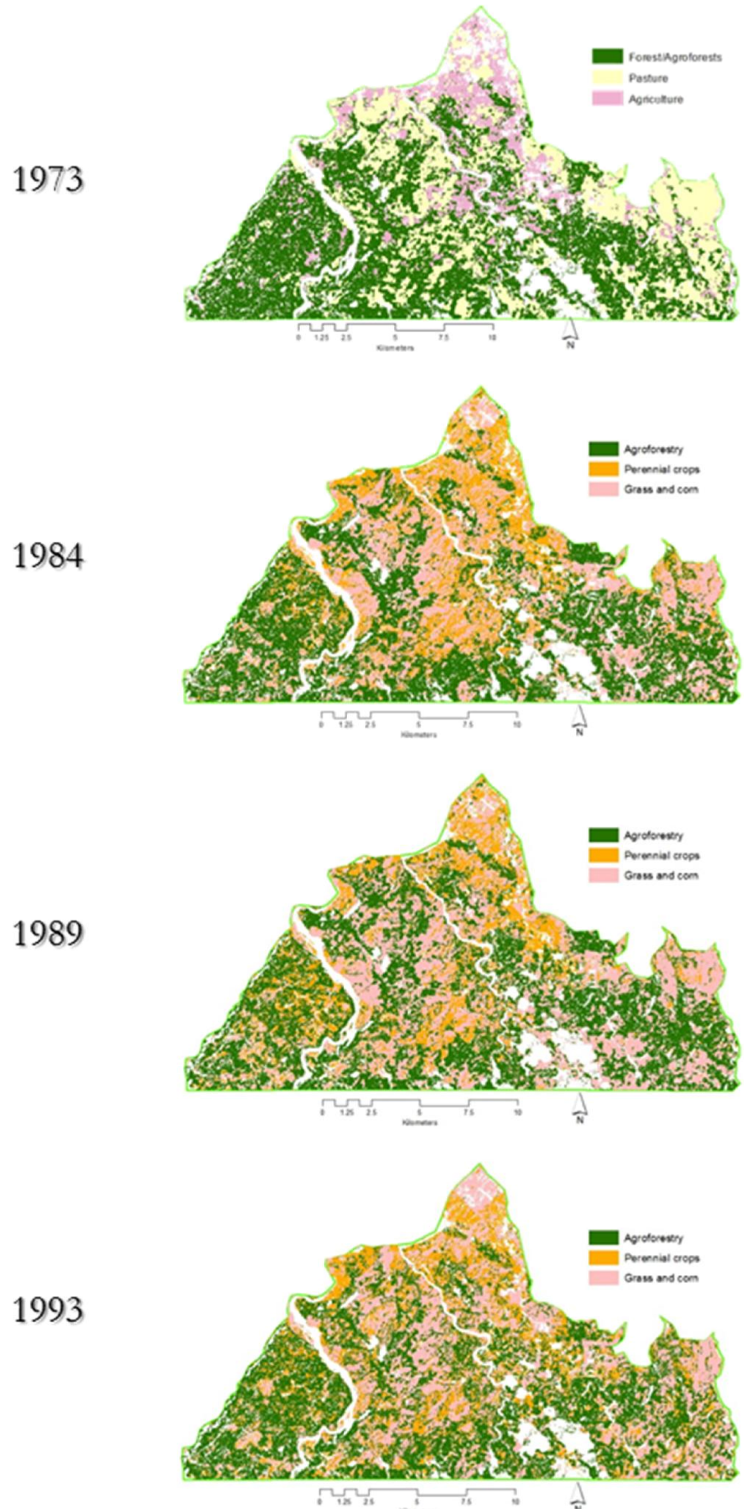


Figure 2.5. Land use maps low land Atzalan 1996-2006

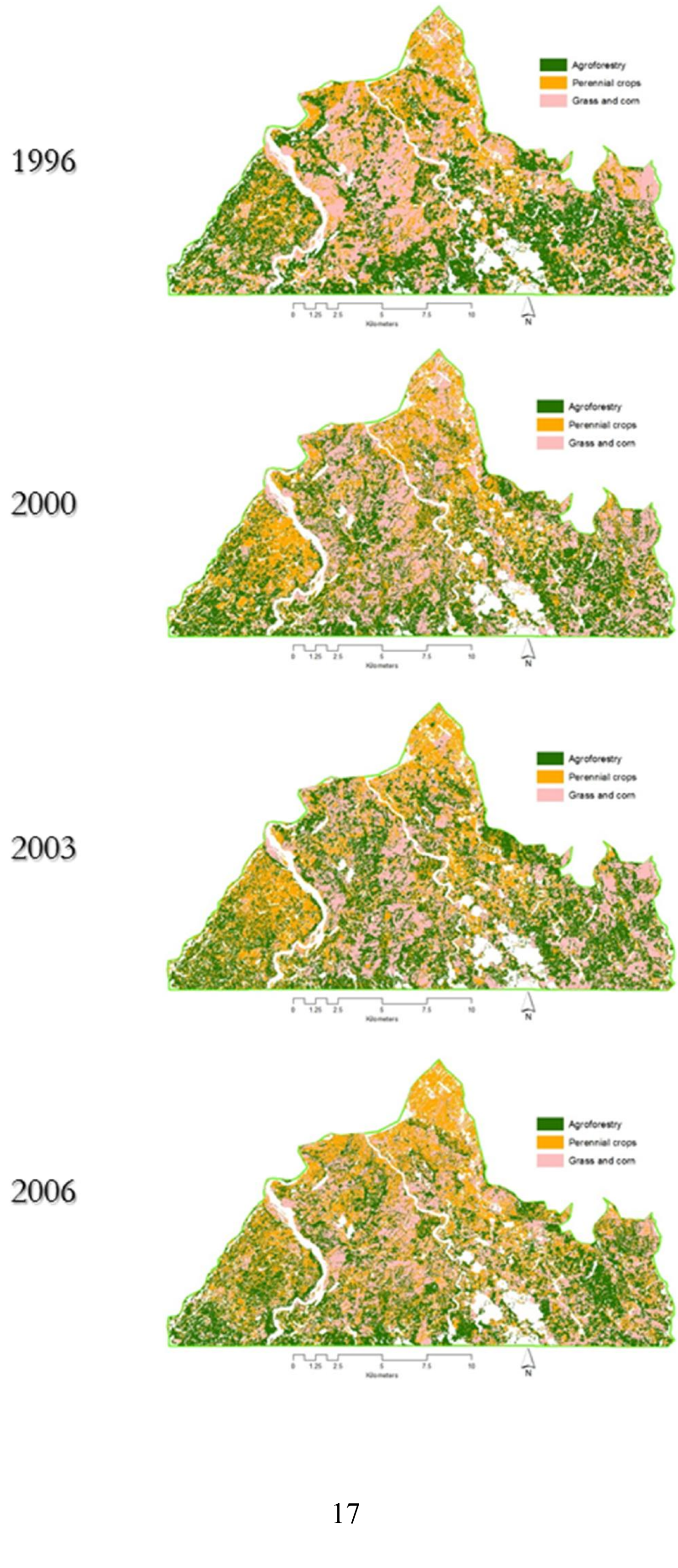
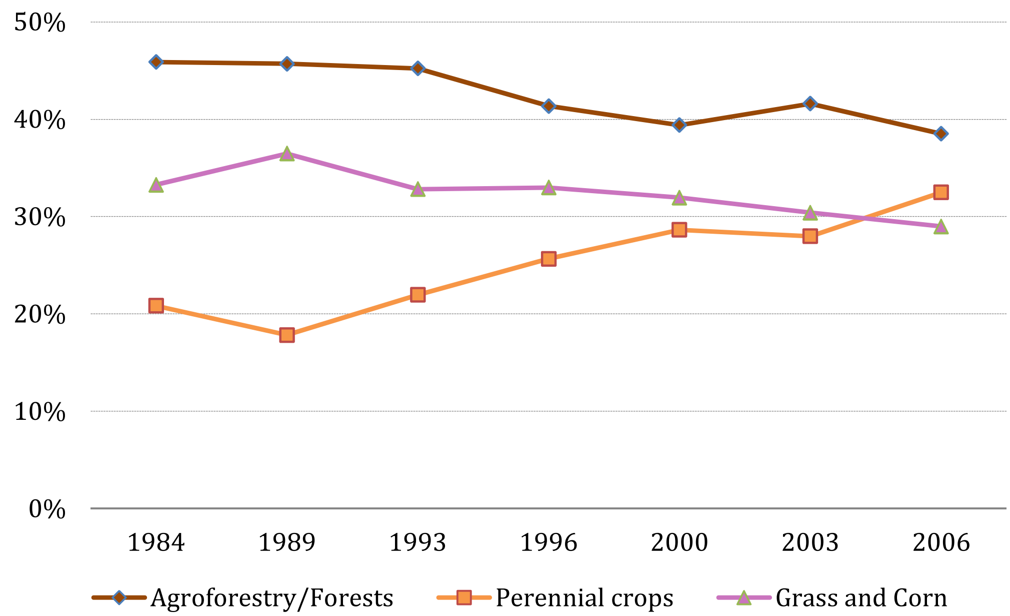


Figure 2.6. Aggregated land use trends 1984 – 2006



Different methodological approaches to account for spatial dependence in the context of discrete choice random utility models are analyzed by Schiern and Felthoven (2011) and by Robertson et al. (2009). A commonly implemented workaround to deal with spatial autocorrelation is to construct spatial correlograms to identify at what distance the spatial autocorrelation ¨vanishes¨, and at that aggregation level implement systematic sampling, or systematic random sampling, to generate a sample of spatially independent observations (Dunn & Harrison, 1993). The Spatial Statistics ArcGIS toolbox was used to estimate the value of the global Moran's I index at different distance values between observations to generate a systematically selected random sample of independent observations that are not spatially clustered at the 5% significance level. In the first part of the sampling process the aggregated and disaggregated land use classifications produced with the 2006 satellite image were used to identify the distance at which the potential sample points were not spatially related. Figure 2.7 shows that at around 720 and 890 meters of separation between sampling points the global Moran's I index is below the threshold level. After repeating the analysis for the rest of the land use maps a separation of around 890 meters between sampling points appeared to be statistically adequate during all the points in time for which we have remotely sensed data. This sampling mechanism produced a sample composed of 274 sampling points distributed across the study area. We follow the common approach in the land use literature and consider that each sampling point correspond to a parcel with a land use value determine by a K-nearest neighbors algorithm with k equal to 25 since there is not available information to identify the polygons corresponding to the area of the parcels in the study region. Figure 2.8 shows the distribution of the sample parcels in the region under analysis.

Figure 2.7. Correlogram of 2006 land use data

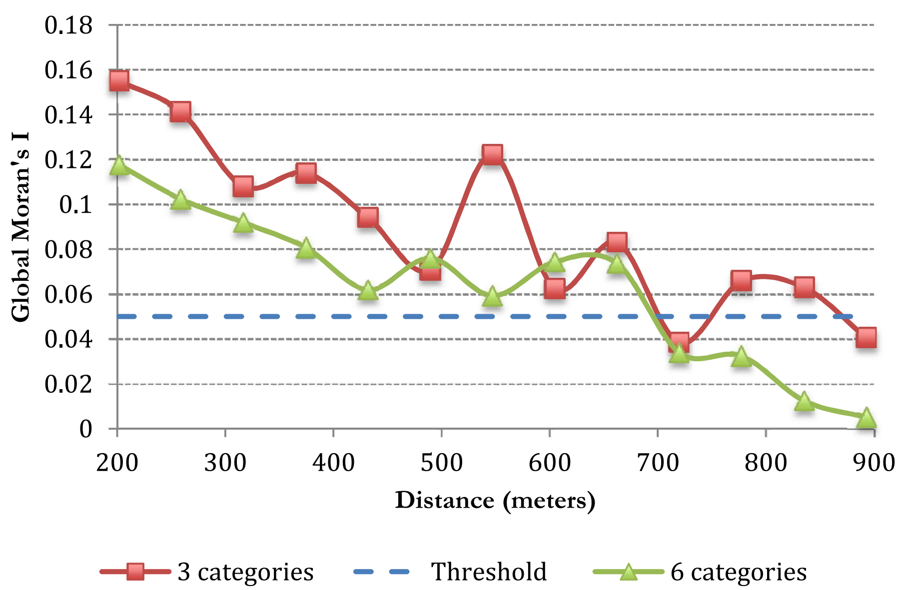
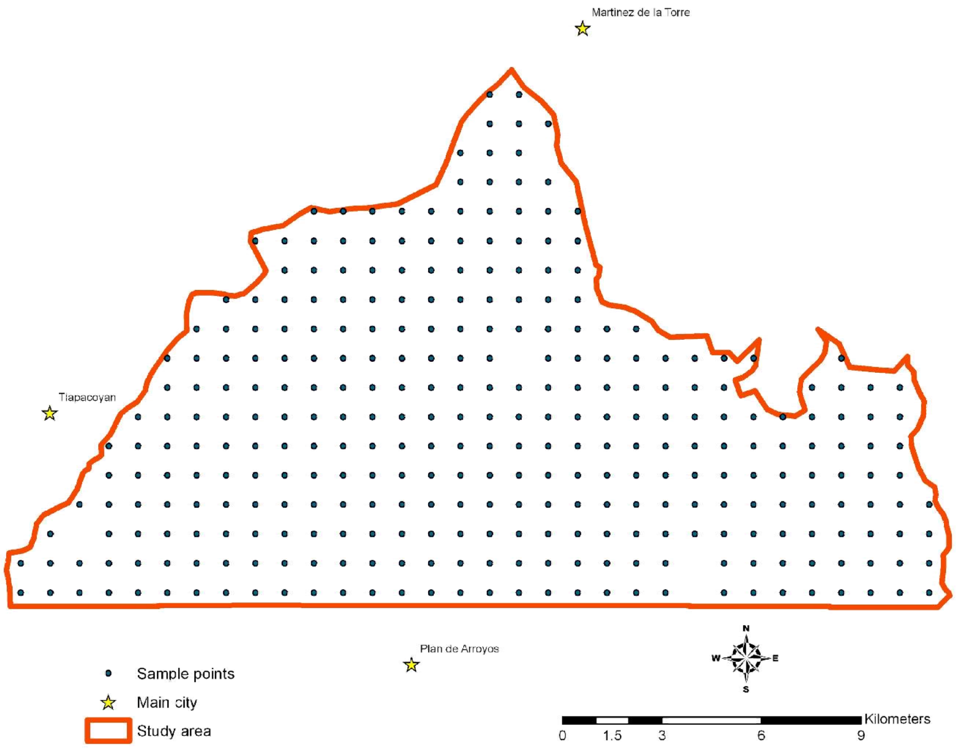
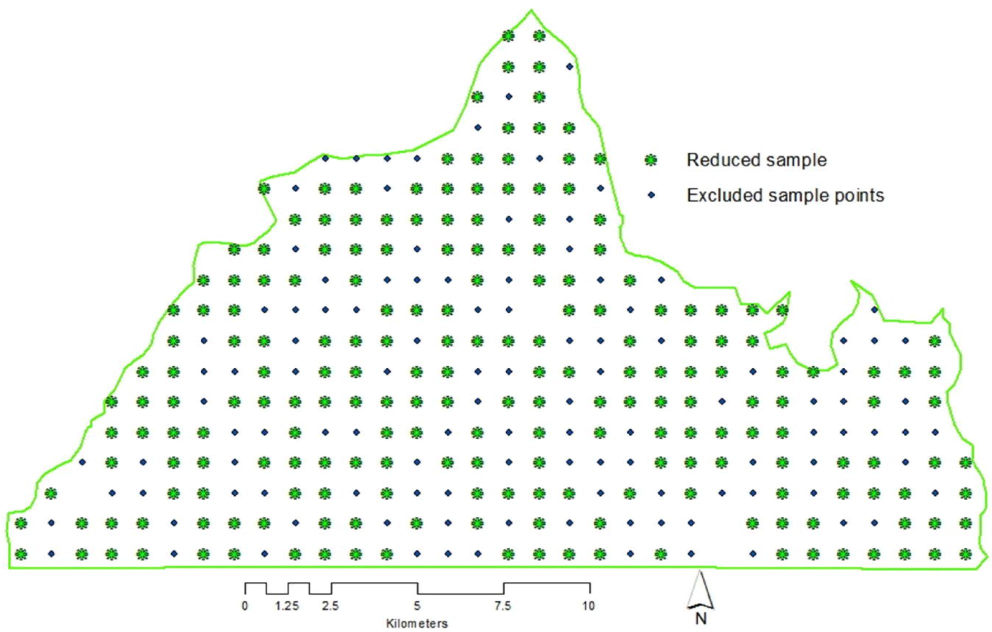


Figure 2.8. Location of the systematically selected sampling points



The sample was further modified to allow the estimation of the age of the land use observed in each parcel, which is a relevant variable in the dynamic model of land use decisions implemented in this thesis. Recall that in the original dataset of land use classifications the largest separation between observations occurs between 1973 and 1984. To define a starting value for the age variable required in the dynamic analysis (more details can be found in chapter 5), in our sample of 274 parcels we identified 210 parcels that did not change land uses during the period 1973 - 1984. In this subset we consider that the minimum age of each land use at the beginning of 1984 is 11 years (i.e., $age_{i,\,1984} = 11$ for all parcels $i = 1,...,210$), which corresponds to mature land uses that have reached its maximum productivity. To be able to compare the result obtained from all the different modeling approaches used in this research we limit our analysis to the subset of 210 parcels and land use data for the periods 1984, 1989, 1993, 1996, 2000, 2003, and 2006. Figure 2.9 shows the location of the observations included in the final sample of 210 parcels.

Figure 2.9. Location of the sample parcels.

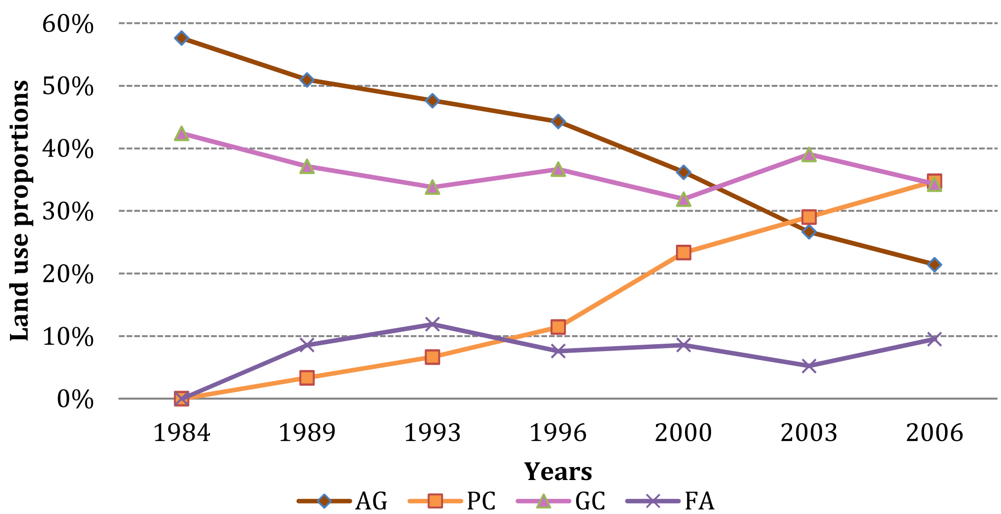


2.2.3. Fallowed lands category

Land use change is in most of the cases a costly action since it requires the removal of the current land use, an up-front investment to establish a new crop, and the financial resources to implement maintenance activities during the growing period of the newly planted crops. Under some circumstances agents would prefer to abandon their lands during some period or periods instead of changing the use of their land. To control for this type of decision, land use information produced with the remotely sensed data was used to construct an additional land use category composed by fallowed parcels

(FA). This land use type was assigned to some parcels after analyzing the sequence of land use decisions produced with the remotely sensed data. I consider that a temporary land use transition that lasts at most six years (roughly two observation intervals) from GC or PC to AG and then back to the previously observed land use indicates that that parcel was in fact fallowed during the period detected as AG. An example may clarify the procedure. Consider that after classifying the remotely sensed data the land use in parcel *s* is classified as GC during 1996, AG during 2000 and again GC in 2003. This land use sequence is not logical either by economic or biological reasoning. In cases like this we consider that parcel *s* was in fact fallowed during 2000 and that the classifier algorithm categorized the land use as AG after detecting an increase in biomass that was likely generated because the landowner forwent maintenance activities in that parcel. This procedure allows me to identify potential fallowed parcels mainly in the GC category since temporary land use transitions between PC and AG represent less than 0.15% of the sample dataset. We do not consider this a relevant drawback of the procedure to detect FA parcels in the PC category since land abandonment of citrus or banana plantations is not common given the potential impacts on future yields that the lack of maintenance activities may have. Figure 2.10 shows the trends across the four land use categories in the sample data during the period 1984 – 2006. In that figure we can observe that the upward trend in the perennial crops class, and the slight decline in the proportion of land devoted to grasslands and cornfields observed in the complete dataset are present in the sample data. On the other hand, the decline in the proportion of agroforest/forest parcels is higher in the sample data than in the whole study region.

Figure 2.10. Land use proportions in the sample data 1984 – 2006



2.2.4. Land use drivers

Reviews of studies in the land use and land cover change literature have found that economic factors, political conditions, technological development, demographic changes, cultural differences and bio-geographical elements interact at different temporal and spatial scales inducing modifications to the terrestrial ecosystem (Bürgi et al., 2004; Geist & Lambin, 2002; Redman et al., 2004). Given that a significant percentage of the land use and land cover change literature has been focused on understanding the causes of deforestation, some general conclusions have been reached about how changes in some of the underlying drivers of landscape dynamics impact deforestation rates. For instance, Angelson and Kaimowitz (1999) after reviewing 146 studies of deforestation in tropical areas conclude that increases or improvements in the road network, increases in agricultural prices, low wages, and low employment opportunities generally increase deforestation. Nevertheless, those authors also find mixed result for other variables such as population pressure, poverty, and income levels. The expected direction of the marginal effects of changes in drivers of land use change becomes less clear when trying to analyze landscape dynamics using a broader set of land use categories.

Additionally, the results may be highly influenced by the methodological approaches, modeling assumptions and the quality of the dataset used in the analysis. For instance, a study considering that market prices drive landowners decisions may find counterintuitive results if decision makers are not fully market integrated and use shadow prices that may be affected by subsistence constraints or by cultural factors to make their land use decisions (Angelsen & Kaimowitz, 1999; Arslan & Taylor, 2008). The next subsections present a description of the drivers of land use change included in this thesis, a justification of their inclusion in the model as well as a discussion of the expected direction of the marginal effects.

2.2.4.1. Market prices

Agricultural prices are one of the most important driving forces of land use change (Angelsen & Kaimowitz, 1999). Typically, in economic models of household land use decisions, agricultural prices are commonly assumed to be exogenously determined in the market. Nevertheless, some studies show that in regions with non-competitive market structures, imperfect information, insecure property rights, high transportation costs, or with non-market values attached to agricultural activities, subjectively determined shadow prices may guide land managers' resource allocations (Angelsen & Kaimowitz, 1999; Brooks, 2010; De Janvry, Fafchamps, & Sadoulet, 1991). Studies that estimate the relationship between land use decisions and fluctuations in market prices in those regions may misleadingly estimate decision makers as insensitive to market prices, or not economically rational, if the shadow prices that they use are significantly different from the market reference (Arslan & Taylor, 2008; Puri, 2006).

As mentioned in previous paragraphs, the study region was selected after two preliminary studies by Ellis et al. (2010) and Baerenklau et al. (2012) found that agents in the low land area of Atzalan, Veracruz, Mexico present a relatively high elasticity of substitution between land uses. Those authors observe that a significant proportion of the agents in that region replaced their coffee farms for citrus or banana plantations in response to low coffee prices. Considering those agents as market-price responsive, I use time series data on average market prices received by farmers at the state level per ton of coffee, lemon, orange, tangerine, mandarin, grapefruit, banana, livestock, and corn obtained from the Secretariat of Agriculture, Livestock, Rural Development, Fisheries and Food (SAGARPA, 2012) to construct price indexes for the land use categories considered in this research.

Since some of the land use categories are integrated by more than one agricultural product it was necessary to merge the price information from each component of the aggregated land use categories into a useful variable. For the AG category, given that there is not commercial use of forested lands in the study region and that the main component of the agroforestry production system is coffee, I use the average rural price per-ton received by coffee growers as representative of the price for this category (see table 2.3 in the Appendix section). Since the PC category is integrated by different citrus varieties as well as banana plantations; to construct a price index for this category it was necessary to implement a two-step procedure. In the first step price information of citrus varieties harvested in the study region was used to construct a weighted average price per-ton, with weights set according to the area harvested for each citrus type at the state level (see tables 2.4 and 2.5 in the Appendix section). In the second step, a similar weighting process was implemented to merge the citrus price index information with the time series data of average prices of banana observed in rural areas of the state of Veracruz (see table 2.6 in the Appendix section).

A different procedure was used to construct the price index corresponding to the GC category. Empirical observations indicate that agricultural activities in the study area are developed using labor and land intensive production technologies that have not been significantly modified in decades. This is particularly true for cornfields and grassland parcels in which it is fair to assume that on average farmers get the same amount of grain and weight gain of livestock per hectare independently of the age of the plantations. Obviously this is arguable since the weight gain of livestock per hectare is dependent on the age of the herd; and the cornfields productivity may decrease through time if the soil is not fertilized or if the landowner does not implement a rotational production system. Additionally, climate factors may induce variability in the productivity of the components of the GC category. Nevertheless, since collecting information about soil quality or fertilization practices at the parcel level or about the age composition of the livestock ranching in grasslands is not feasible for this research, I use the average productivity of corn plantations (that is estimated to be 2.31 tons of corn per hectare) (SAGARPA, 2012); and the average livestock weight gain per hectare observed in unfertilized grasslands in the state of Veracruz, Mexico (that is 425 kilograms per year, around 937 pounds), (Tergas & Sanchez, 1979) to construct a per hectare weighted price index for the GC category. The information used to construct this index is presented in table 2.7 in the appendix.

Since the FA category does not involve crop production, to account for the monetary reward that a farmer that decides to let his land fallow can get in an alternative activity I use the yearly minimum wage for construction workers as a price index for this category (see table 2.3 in the Appendix section). Farmers in the state of Veracruz do not have many employment options. Besides working in parcels owned by other people, the most common option in the study area is to migrate to Mexico City or to the U.S.A. (Nava-Tablada & Martínez-Camarillo, 2012). Given the educational level of farmers in the study area the salary received by construction workers can be used as proxy of the income that farmers can get working off parcel.

The procedures followed to construct price indexes for the land use categories used in the analysis generate an average price per ton for the AG and PC categories, an average price per hectare for the GC category, and an average price per farmer for the FA category. The differences in the procedures are only to facilitate numerical analysis, and the units are reconciled by using yield values that produce consistent per hectare revenue estimates.

2.2.4.2. Yields

Historical data from the Agro-food information system (SAGARPA, 2012), and information from agronomists working in the region is used to determine the average productivity per hectare of shade grown coffee, banana, citrus, pasture, and corn. That information is used to estimate expected yields for AG, and PC plantations for ages ranging from one to twenty-five years (table 2.1). On the other hand, since the GC price index is per se a measure of the expected revenue per hectare that is independent of the age of the land use, a unit value is assigned through all the aforementioned age range.

Table 2.1. Expected yield per hectare at different plantation ages for the AG and PC categories and yield indexes for the GC and FA categories.

Age	AF	PC	GC	FA	Age	AF	PC	GC	FA
1	0.00	0.33	1	0.5	14	2.24	12.25	1	0.5
2	0.00	1.29	1	0.5	15	2.24	12.25	1	0.5
3	0.35	1.29	1	0.5	16	2.24	12.25	1	0.5
4	0.80	1.29	1	0.5	17	2.24	12.25	1	0.5
5	1.40	4.85	1	0.5	18	2.24	12.25	1	0.5
6	2.10	8.15	1	0.5	19	2.24	12.25	1	0.5
7	2.24	10.02	1	0.5	20	2.24	12.25	1	0.5
8	2.24	12.25	1	0.5	21	2.24	12.25	1	0.5
9	2.24	12.25	1	0.5	22	2.24	12.25	1	0.5
10	2.24	12.25	1	0.5	23	2.24	12.25	1	0.5
11	2.24	12.25	1	0.5	24	2.24	12.25	1	0.5
12	2.24	12.25	1	0.5	25	2.24	12.25	1	0.5
13	2.24	12.25	1	0.5					

The AG and PC yield units are tons. The GC and FA yields are indicators that map the corresponding price index to the corresponding proportion of the annual revenue per hectare that is independent of the age of the plantation.

For the FA category, the constructed price index represents the revenue that a farmer can get by working off-farm letting his land fallow. Empirical observations in the study area indicate that one person can complete all the required maintenance activities for a 2-hectare parcel without needing to hire additional labor. Therefore, I assume that the price index of the fallow category represents a unit of "yield" for a two-hectare parcel; or alternatively that the FA revenue that can be linked to a one hectare parcel corresponds to a yield equivalent of 0.5 of the FA price index. Again all these differences in the construction of the price indexes are for convenience in the numerical analysis. Alternative approaches can be implemented but the estimation of the revenue per hectare will reach similar values.

2.2.4.3. Parcel specific characteristics

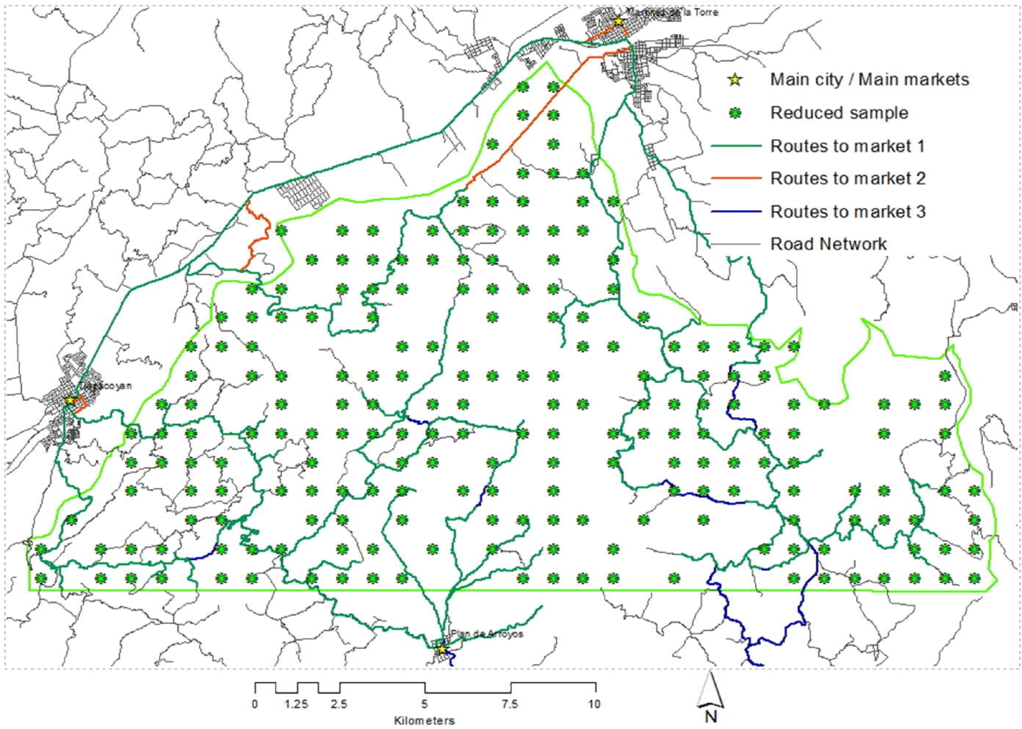
Since this research is based on Von Thünen's model of land use that assumes that land is allocated to the use with the highest return, and that the revenue and cost associated with each land use option are determined by spatially explicit variables (Angelsen, 2007; Chomitz & Gray, 1996), I include in the dataset parcel specific characteristics that are usually included in models of landscape dynamics.

2.2.4.3.1. Distance to nearest road and to nearest market

In spatially explicit random utility land use models it is commonly assumed that the distance from a parcel to a market affects the revenues and costs associated with the land uses available to the decision makers. It is logically assumed that the farther a parcel is to a road or to a market, the higher the transportation costs to sell the harvest or to purchase the necessary inputs to cultivate a farm. Studies of deforestation processes have found that the closer a parcel is to a road the higher the likelihood of that area to be deforested, although in some cases the road infrastructure is an endogenous variable defined by decision makers to provide access to forests stands with high quality timber (Angelsen & Kaimowitz, 1999; Chomitz & Gray, 1996). In shade-grown agroforestry regions the proximity to markets may reduce the probability of substitution of agroforestry areas for other land use types (Blackman, Ávalos-Sartorio, & Chow, 2012). On the other hand, in a study of land use change in shade-grown coffee based agroforests in Mexico developed by Blackman et al. (2008) the authors find that parcels that are farther from markets and big cities are more likely to be replaced by subsistence agriculture. Following the last two mentioned studies, I expect a reduction in the probability of observing cash crops (AG or PC) in the region as the distance to a market increases, and an increase in the likelihood of agents selecting subsistence crops (such as corn), land uses that require a large area (such a cattle ranching activities), or land abandonment as the distance to a marker increases.

There are three main regional market centers in the proximity of the study region at which farmers can sell their products. Those three markets have similar prices for the produce generated from the land use categories under analysis. To compute the distance from each parcel to the nearest market I followed a three-stage process. First, I computed the Euclidean distance from each sample parcel to the nearest road using vector data from the National Institute of Geography and Informatics (INEGI, 1999). Second, by using the network analysis ArcGIS extension and vector data of the road network in the area, I computed the most efficient route (in terms of distance) from the nearest road (corresponding to each sampling point) to every market center. Finally, I compared the distances to each market and selected the shortest one. This variable is considered to be static since the road network was not significantly expanded during the period of analysis. It is true that improvements were made to the conditions of some of the main roads (e.g., changing from dirt roads to paved roads) which potentially reduced diving time but not driving distance to each market. Figure 2.11 shows the location of the relevant cities identified as main markets as well as the road network in the study area.

Figure 2.11. Road network in the study area.



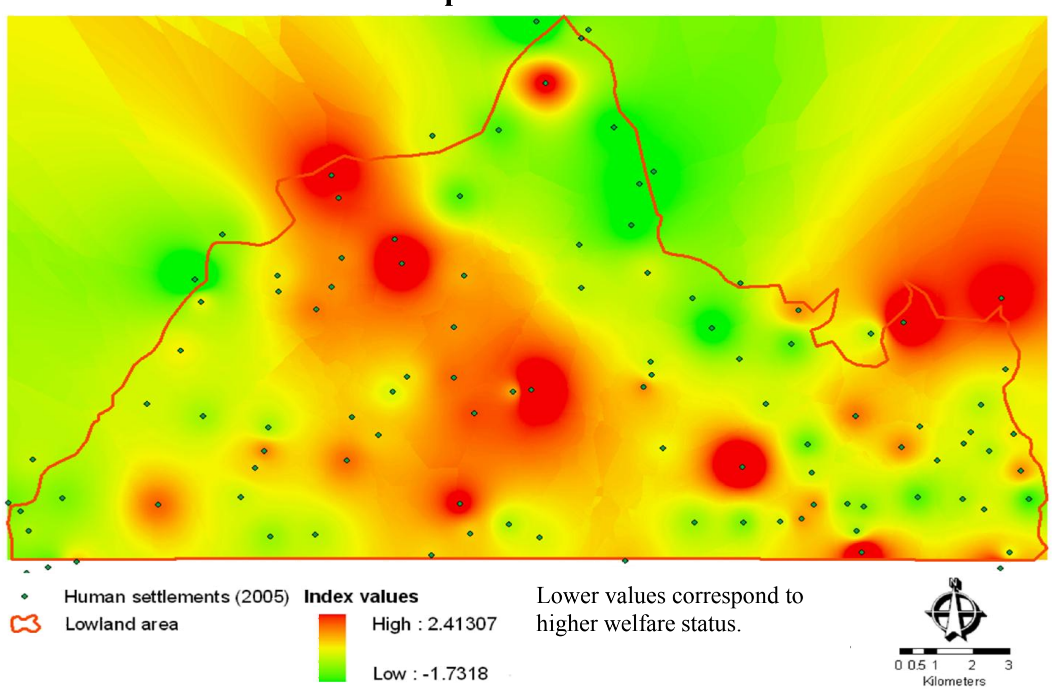
Note: The routes to the different markets overlap in some road segments. The displayed routes do not necessarily correspond to the most efficient ones.

2.2.4.3.2. Poverty index

We can argue that agents in poor areas relying on household labor to implement agricultural activities may be less likely to stop cultivating their parcels (Albers, Avalos-Sartorio, Batz, & Blackman, 2006). On the other hand, farmers in a region with better welfare status may require relatively higher salaries to work in agricultural activities potentially increasing the cost of labor and the probability of observing fallowed parcels. In the agricultural region under analysis I expect that areas with lower poverty rates are likely to be composed of PC or AG (i.e., by cash crops that require higher up-front investment) than subsistence crops.

Starting in 1995 the Mexican Government computes every 5 years an index that uses data regarding education accessibility, housing conditions and monetary income to measure the degree of poverty at the community level. This index in general ranges from -2.37 to 4.49, with lower values corresponding to a better welfare status (CONAPO, 2006). A review of the statistics generated by CONAPO (1998, 2006, 2011) indicates that the poverty level in the 104 communities located either within the study area or up to 500 meters outside its boundary, has not fluctuated significantly during the 1995, 2000, 2005 and 2010 analyses. We use the Inverse distance weighting method (IDW) to interpolate the 2005 poverty index values across the study area and the results with the corresponding range of the poverty index are shown in figure 2.12.

Figure 2.12. Poverty index computed using the Inverse Distance Weighted interpolation method.



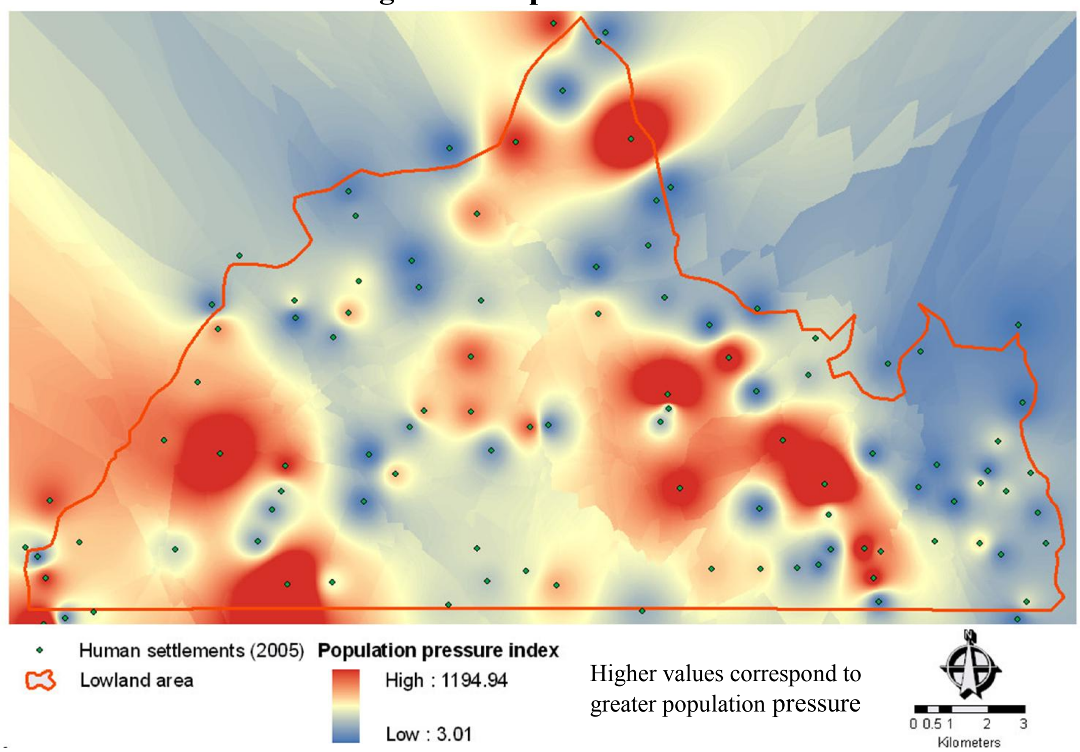
2.2.4.3.3. Population density

Several studies have documented the influence that population growth has over land use change in rural areas (see for instance Andersen, 1996; Meyer & Turner B L, 1992; Verburg, Veldkamp, & Bouma, 1999). To capture the effects of population density and labor availability on land use change for the parcels contained in the sample we use population data at the community level and interpolate the information across the study area. Statistics from CONAPO (1998, 2006, 2011) indicate that the number of inhabitants in most of the communities has not significantly changed during the study window. It is expected that large human settlements generate more pressure over their surrounding environment and at the same time provide more labor to harvest the land. Since both effects diminish as the distance to the settlement increases I use population data from 2005 and the inverse distance weighting interpolation procedure to construct the interpolated surface shown in figure 2.13.

2.2.4.3.4. Slope, elevation, and soil texture

Topographic variables can constitute a relevant factor in the configuration of the choice set of land uses available to the decision maker or alter crop productivity levels. For instance, soil chemical and physical properties that impact soil quality can be dependent on the degree of slope and on the land use/land cover implemented in a particular area (Moges & Holden, 2008; Wei, Fu, Horton, & Shao, 2010). Additionally, the land use options for a parcel may be constrained by the degree of slope and soil type observed in that piece of land (Bakker et al., 2005; Fu et al., 2006). Furthermore, parcels with high slope values may be prone to soil degradation if the vegetation in an area is dramatically changed (Lal, 2001; Van de Koppel, Rietkerk, & Weissing, 1997). In the study region empirical observations indicate that areas with high degree of slope are used mostly for agroforestry production while parcels with low slope are preferred for cornfields, grasslands, citrus or banana production.

Figure 2.13. Population pressure index computed using the Inverse Distance Weighted interpolation method.



On the other hand, microclimatic conditions that depend on altitudinal gradients may also constrain land use options. Agro-ecological requirements for the production of citrus, coffee and banana suggest that parcels at higher elevations may be more likely to be devoted to agroforestry production, citrus instead are mostly found at lower elevations. Since corn and grass can be produced in parcels located at different elevation gradients, the direction of the marginal effects could go in either direction depending on the values of other land use drivers.

To study the effect of topographic variables in the land use decision making process followed by agents in the geographical region under analysis, I use vector data of elevation level curves obtained from INEGI (1998) to construct a digital elevation model that was used to generate slope and elevation information. Figures 2.14 and 2.15 show the raster data of slope and elevation. Additionally, soil texture information from SEMARNAP (1998) was used as a proxy of soil quality although there is not significant variation across the study region (figure 2.16) .

Figure 2.14. Map of slope information.

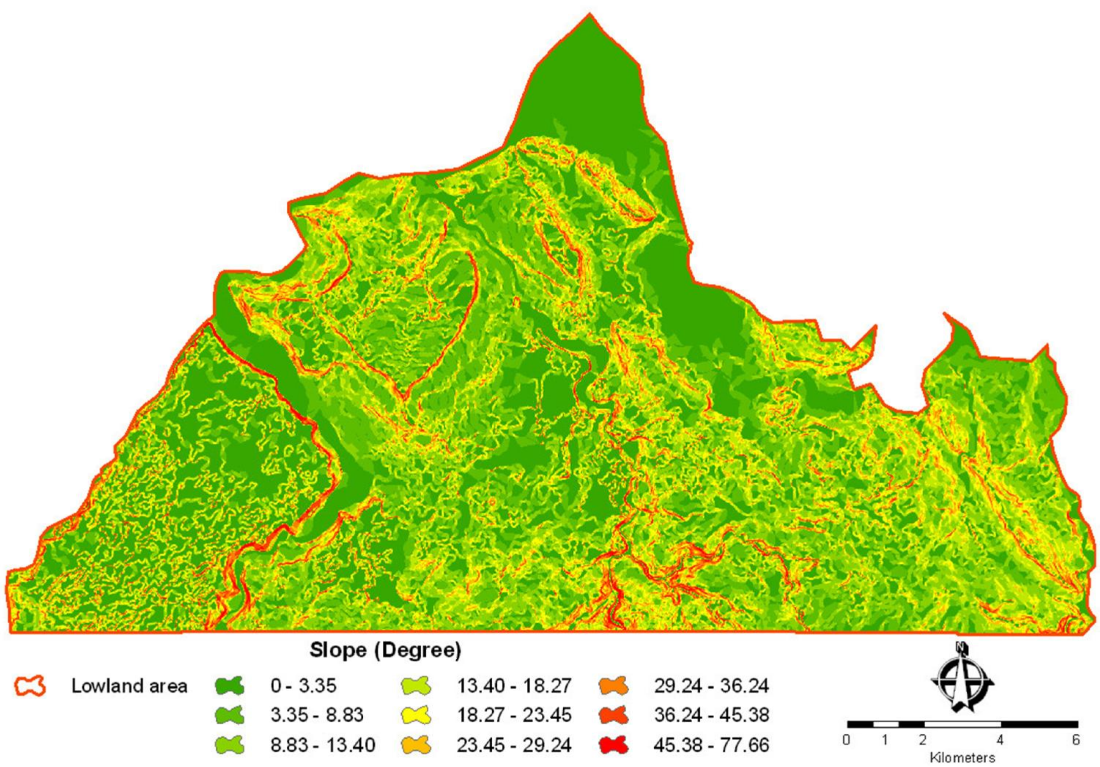


Figure 2.15. Map of elevation.

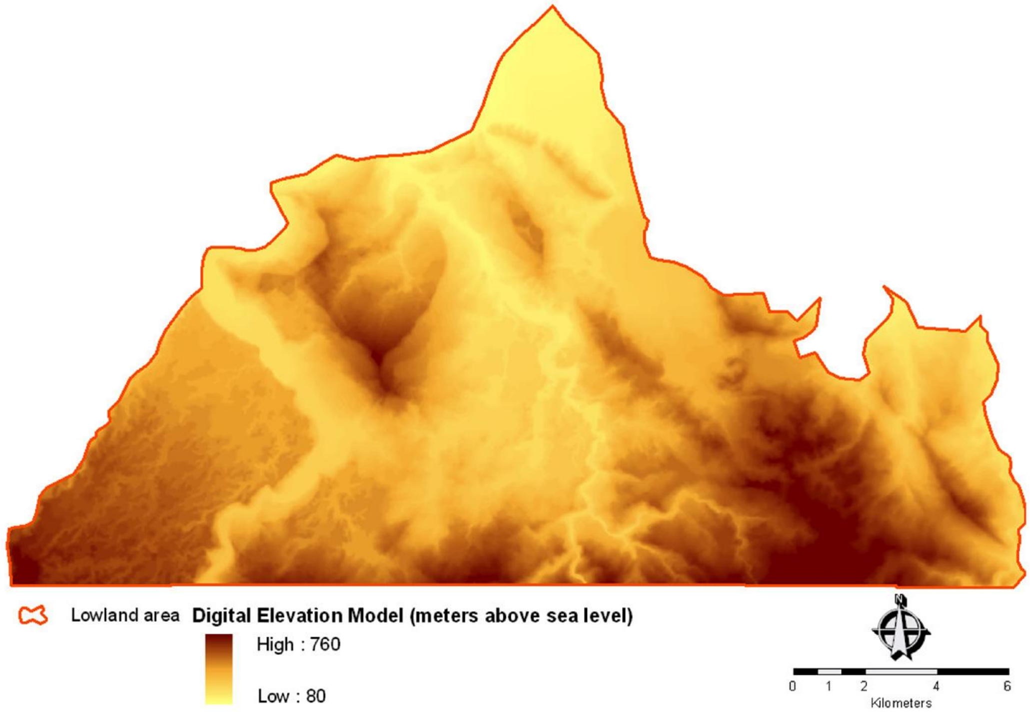
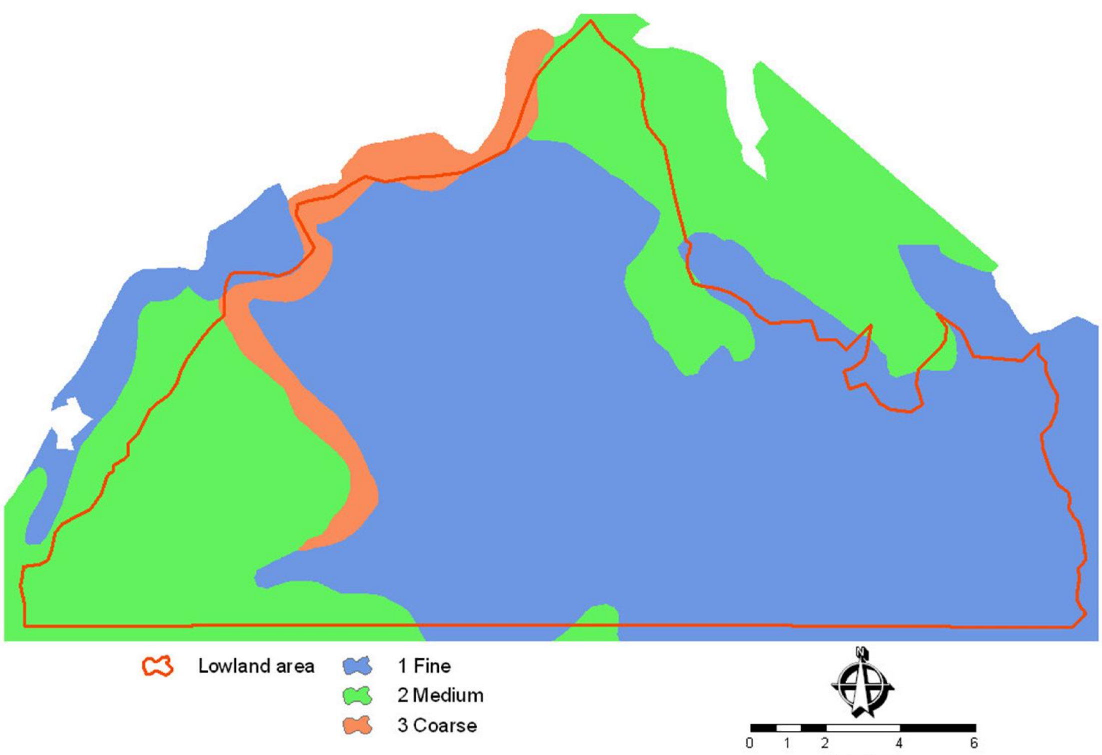


Figure 2.16. Map of soil texture.



To conclude this chapter Table 2.2 presents a summary of the mean, minimum and maximum values of the parcel specific characteristics in the sample data used in the analysis. Further details are provided in the remaining chapters of this thesis.

Table 2.2. Summary statistics for the parcel specific variables

Variable	Description	Mean	Min	Max
Elevation	Meters above sea level	354	85	726
Slope	Degrees	10.49	0	60.09
Poverty	Index that uses education accessibility, housing conditions and monetary income data to measure the degree of poverty with lower values corresponding to a better welfare status	0.316	-0.798	2.109
Population	Index to measure labor availability	263	30	793
Soil texture	Soil texture of parcel (1 = fine, 2 = medium, 3 = coarse)	1.34	1.00	3.00
Distance to road	Euclidean distance from each parcel to the nearest road (m)	389	0	1,779
Distance to nearest market	Distance from each parcel to nearest market (km)	14.36	2.93	35.52

Summary statistics are computed for all parcels and for all the time periods considered in the analysis.

2.3 Appendix.

Table 2.3. AG and FA price indexes.

	Nominal values			Real values	
	AG	FC		AG price index	FA price index
	Average rural price of cherry coffee (Mex $ / Ton)	Construction workers' salary (Mex $ per year)	IPC (2000)	Average rural price of cherry coffee (2000 Mex $ / Ton)	Construction workers' salary (2000 Mex $ per year)
1980	9.25	68.54	0.0012	7500	55,578
1981	9.90	88.42	0.0016	6275	56,038
1982	12.00	123.67	0.0025	4786	49,324
1983	15.00	206.83	0.0051	2963	40,863
1984	33.00	317.03	0.0084	3941	37,857
1985	59.32	491.70	0.0132	4490	37,220
1986	230.00	830.45	0.0246	9349	33,755
1987	380.00	1770.12	0.0570	6663	31,036
1988	550.00	3347.52	0.1221	4503	27,406
1989	661.39	3774.84	0.1466	4512	25,752
1990	435.00	4334.94	0.1857	2343	23,349
1991	594.48	5004.00	0.2277	2610	21,974
1992	573.00	5605.92	0.2630	2178	21,312
1993	586.18	6001.92	0.2887	2030	20,790
1994	850.00	6422.40	0.3088	2753	20,798
1995	2139.00	7553.76	0.4169	5131	18,120
1996	2439.95	9368.40	0.5602	4355	16,723
1997	3294.65	12672.00	0.6757	4876	18,753
1998	4395.57	14443.20	0.7834	5611	18,437
1999	4838.16	14443.20	0.9133	5297	15,814
2000	2855.43	15883.20	1.0000	2855	15,883
2001	1472.92	16920.00	1.0637	1385	15,907
2002	1153.69	17683.20	1.1172	1033	15,828
2003	1373.42	18316.80	1.1680	1176	15,682
2004	1619.62	18984.96	1.2227	1325	15,527
2005	2177.47	19641.60	1.2715	1713	15,447
2006	3055.76	20427.84	1.3177	2319	15,503
2007	2845.69	21225.60	1.3699	2077	15,494
2008	4016.67	22075.20	1.4401	2789	15,329
2009	4133.68	23002.56	1.5164	2726	15,169
2010	4630.81	24001.92	1.5844	2923	15,149
2011	5496.29	26159.04	1.6333	3365	16,016

Source: Author calculations based on data from SAGARPA (2012) and INEGI (2009).

Table 2.4. Nominal average market prices for relevant citrus varieties harvested in the study region.

	Surface harvested (Ha.)					Average rural price (Mex $ / Ton.)				
Year	**Lemon**	**Mandarin**	**Orange**	**Tangerine**	**Grapefruit**	**Lemon**	**Mandarin**	**Orange**	**Tangerine**	**Grapefruit**
1980	5,728	5,039	71,683	0	3,434	3	2	4		6
1981	7,935	6,095	72,427	0	3,434	4	3	4		6
1982	15,281	13,465	86,752	0	3,455	9	3	7		7
1983	15,486	13,425	86,861	0	3,435	18	20	16		21
1984	16,014	13,579	87,939	0	3,434	34	28	25		22
1985	6,756	6,540	85,713	0	1,233	22	66	21		16
1986	6,527	6,540	88,710	0	3,300	30	76	26		23
1987	5,317	5,389	90,557	0	4,462	133	115	50		80
1988	5,339	4,492	94,617	0	4,462	400	187	170		190
1989	6,932	4,492	120,291	0	4,462	844	223	178		203
1990	6,216	4,492	120,291	0	4,462	1,090	279	357		250
1991	7,176	4,727	114,115	0	4,493	984	286	423		292
1992	10,266	4,842	120,975	5,788	4,200	700	600	450	319	471
1993	10,537	5,631	123,216	5,452	4,066	553	570	456	358	697
1994	11,284	5,154	133,473	6,381	4,204	550	550	168	400	1,200
1995	11,821	5,697	153,048	5,452	4,204	1,110	356	577	420	641
1996	10,966	5,511	155,093	5,452	4,174	1,878	769	547	349	361
1997	10,894	5,739	144,909	5,452	4,276	1,577	566	541	400	489
1998	12,704	7,331	152,707	10,853	3,654	1,689	1,007	623	2,000	1,273
1999	13,356	5,722	144,418	10,050	3,494	1,964	1,397	1,001	2,009	1,278
2000	13,965	6,181	144,082	10,695	4,283	740	1,092	764	1,000	1,880
2001	16,589	6,601	147,016	11,352	4,306	1,482	504	483	786	625
2002	18,499	7,096	153,981	13,230	4,638	1,466	904	589	913	1,693
2003	20,820	7,477	148,176	12,930	5,353	1,430	884	741	853	887
2004	21,990	7,596	150,758	12,655	5,622	1,471	500	674	970	768
2005	22,177	7,519	145,513	12,393	5,540	1,610	771	488	1,116	1,646
2006	23,009	7,674	144,613	12,659	5,380	1,728	1,036	663	1,297	972
2007	27,374	7,826	152,395	13,154	5,447	1,308	932	762	1,018	1,377
2008	31,210	8,576	160,411	14,663	5,662	1,732	826	707	796	790
2009	31,245	8,106	157,117	14,515	5,634	2,047	683	748	1,383	819
2010	31,300	8,137	157,798	14,541	6,221	2,574	1,280	999	1,254	1,115
2011	33,003	8,944	157,586	11,618	5,674	3,039	843	1,225	1,898	1,228

Source: Author calculations based on data from SAGARPA (2012).

Table 2.5. Weights used to construct a price index for the citrus category.

	Proportions of land devoted to the production of					Proportional prices (Mex $ / Ton.)					(Mex $/ Ton.)
Years	Lemon	Mandarin	Orange	Tangerine	Grapefruit	Lemon	Mandarin	Orange	Tangerine	Grapefruit	Citrus price index
1980	0.067	0.059	0.835	0.000	0.040	0.21	0.13	3.34	0.00	0.22	3.90
1981	0.088	0.068	0.806	0.000	0.038	0.32	0.18	3.55	0.00	0.23	4.27
1982	0.128	0.113	0.729	0.000	0.029	1.21	0.38	4.99	0.00	0.21	6.78
1983	0.130	0.113	0.729	0.000	0.029	2.37	2.24	11.31	0.00	0.61	16.53
1984	0.132	0.112	0.727	0.000	0.028	4.56	3.17	18.13	0.00	0.62	26.48
1985	0.067	0.065	0.855	0.000	0.012	1.48	4.28	17.82	0.00	0.19	23.77
1986	0.062	0.062	0.844	0.000	0.031	1.86	4.75	22.30	0.00	0.73	29.65
1987	0.050	0.051	0.857	0.000	0.042	6.70	5.84	42.83	0.00	3.38	58.74
1988	0.049	0.041	0.869	0.000	0.041	19.61	7.71	147.69	0.00	7.78	182.80
1989	0.051	0.033	0.883	0.000	0.033	42.94	7.36	157.54	0.00	6.66	214.49
1990	0.046	0.033	0.888	0.000	0.033	50.03	9.25	317.20	0.00	8.23	384.71
1991	0.055	0.036	0.874	0.000	0.034	54.12	10.36	369.44	0.00	10.04	443.96
1992	0.070	0.033	0.828	0.040	0.029	49.20	19.89	372.69	12.64	13.54	467.96
1993	0.071	0.038	0.827	0.037	0.027	39.13	21.56	377.22	13.11	19.03	470.05
1994	0.070	0.032	0.832	0.040	0.026	38.67	17.66	139.71	15.90	31.43	243.38
1995	0.066	0.032	0.849	0.030	0.023	72.81	11.25	490.00	12.71	14.94	601.71
1996	0.061	0.030	0.856	0.030	0.023	113.65	23.38	467.87	10.49	8.32	623.71
1997	0.064	0.034	0.846	0.032	0.025	100.33	18.96	457.50	12.73	12.21	601.74
1998	0.068	0.039	0.816	0.058	0.020	114.61	39.41	508.27	115.92	24.83	803.05
1999	0.075	0.032	0.816	0.057	0.020	148.13	45.13	816.86	114.06	25.22	1149.40
2000	0.078	0.034	0.804	0.060	0.024	57.68	37.67	614.09	59.68	44.93	814.05
2001	0.089	0.036	0.791	0.061	0.023	132.28	17.91	381.83	48.03	14.49	594.53
2002	0.094	0.036	0.780	0.067	0.023	137.35	32.50	459.31	61.16	39.76	730.08
2003	0.107	0.038	0.761	0.066	0.027	152.86	33.94	563.77	56.61	24.39	831.57
2004	0.111	0.038	0.759	0.064	0.028	162.85	19.11	511.63	61.78	21.73	777.11
2005	0.115	0.039	0.753	0.064	0.029	184.89	30.00	367.37	71.59	47.22	701.07
2006	0.119	0.040	0.748	0.065	0.028	205.62	41.12	496.15	84.94	27.05	854.89
2007	0.133	0.038	0.739	0.064	0.026	173.62	35.37	562.82	64.95	36.37	873.13
2008	0.142	0.039	0.727	0.066	0.026	245.14	32.13	514.52	52.91	20.30	865.01
2009	0.144	0.037	0.725	0.067	0.026	295.28	25.57	542.74	92.68	21.29	977.57
2010	0.144	0.037	0.724	0.067	0.029	369.61	47.78	723.11	83.66	31.82	1255.98
2011	0.152	0.041	0.727	0.054	0.026	462.49	34.79	890.41	101.67	32.13	1521.49

Source: Author calculations based on data from SAGARPA (2012) and INEGI (2009).

Table 2.6. Weights used to construct the PC price index.

	Surface harvested (Ha.)		Precio Medio Rural (Mex $ / Ton.)		Proportions of land devoted to		Price proportions		PC Price index	Consumer Price Index (2000 Mex $)	PC Price index (2000 Mex $)
Year	Citrus	Banana	Citrus price index	Banana	Citrus	Banana	Citrus	Banana			
1980	85884	17478	3.90	1.99	0.83	0.17	3.24	0.34	3.58	0.0012	2901.99
1981	89891	17490	4.27	3.10	0.84	0.16	3.58	0.50	4.08	0.0016	2586.16
1982	118953	17995	6.78	5.20	0.87	0.13	5.89	0.68	6.57	0.0025	2620.58
1983	119207	17966	16.53	19.93	0.87	0.13	14.37	2.61	16.98	0.0051	3354.24
1984	120966	17989	26.48	21.00	0.87	0.13	23.05	2.72	25.77	0.0084	3077.26
1985	100242	21063	23.77	13.35	0.83	0.17	19.64	2.32	21.96	0.0132	1662.25
1986	105077	20703	29.65	20.89	0.84	0.16	24.77	3.44	28.21	0.0246	1146.61
1987	105725	23138	58.74	67.50	0.82	0.18	48.19	12.12	60.31	0.0570	1057.44
1988	108910	26619	182.80	240.00	0.80	0.20	146.89	47.14	194.03	0.1221	1588.50
1989	136177	19995	214.49	421.71	0.87	0.13	187.03	53.99	241.02	0.1466	1644.24
1990	135461	16403	384.71	485.00	0.89	0.11	343.16	52.39	395.55	0.1857	2130.55
1991	130511	16725	443.96	610.75	0.89	0.11	393.53	69.38	462.91	0.2277	2032.71
1992	146071	14280	467.96	550.00	0.91	0.09	426.28	48.98	475.26	0.2630	1806.78
1993	148902	14723	470.05	571.45	0.91	0.09	427.76	51.42	479.18	0.2887	1659.80
1994	160496	14710	243.38	476.00	0.92	0.08	222.95	39.96	262.91	0.3088	851.38
1995	180222	14252	601.71	808.00	0.93	0.07	557.61	59.21	616.82	0.4169	1479.61
1996	181196	14519	623.71	798.22	0.93	0.07	577.44	59.22	636.66	0.5602	1136.47
1997	171270	13902	601.74	504.01	0.92	0.08	556.56	37.84	594.40	0.6757	879.62
1998	187249	13466	803.05	969.17	0.93	0.07	749.17	65.02	814.19	0.7834	1039.33
1999	177040	11793	1149.40	1202.32	0.94	0.06	1077.62	75.09	1152.71	0.9133	1262.12
2000	179204	11278	814.05	1076.70	0.94	0.06	765.85	63.75	829.60	1.0000	829.60
2001	185864	11544	594.53	1014.96	0.94	0.06	559.76	59.35	619.12	1.0637	582.05
2002	197444	11262	730.08	883.63	0.95	0.05	690.69	47.68	738.37	1.1172	660.92
2003	194756	11618	831.57	968.90	0.94	0.06	784.75	54.55	839.30	1.1680	718.59
2004	198622	11830	777.11	1449.80	0.94	0.06	733.42	81.50	814.92	1.2227	666.47
2005	193142	13475	701.07	1784.20	0.93	0.07	655.34	116.36	771.71	1.2715	606.92
2006	193334	13911	854.89	1339.68	0.93	0.07	797.50	89.92	887.43	1.3177	673.49
2007	206195	16393	873.13	2483.74	0.93	0.07	808.83	182.92	991.75	1.3699	723.94
2008	220522	15093	865.01	2410.22	0.94	0.06	809.60	154.39	963.99	1.4401	669.37
2009	216617	14490	977.57	1995.78	0.94	0.06	916.28	125.13	1041.40	1.5164	686.75
2010	217997	14797	1255.98	1963.05	0.94	0.06	1176.15	124.77	1300.92	1.5844	821.09
2011	216824	14867	1521.49	2500.59	0.94	0.06	1423.86	160.46	1584.32	1.6333	970.03

Source: Author calculations based on data from SAGARPA (2012)

Table 2.7. GC Price index.

Years	Nominal values: Corn (Mex $/Ton.)	Nominal values: Livestock (Mex $/Ton.)	IPC (2000)	Real values: Corn (Mex $/Ton.)	Real values: Livestock (Mex $/Ton.)	Real values: Corn (Mex $/Ha.)	Real values: Livestock (Mex $/Ha.)	Real values: GC Price index (Mex $ / Ha.)
1980	5.43	31	0.0012	4403	25136	10171	10683	10427
1981	6.38	37	0.0016	4044	23451	9341	9966	9654
1982	9.14	73	0.0025	3645	29115	8421	12374	10397
1983	19.9	97	0.0051	3932	19164	9082	8145	8613
1984	34.14	179	0.0084	4077	21375	9417	9084	9251
1985	50.55	297	0.0132	3826	22482	8839	9555	9197
1986	89.44	388	0.0246	3635	15771	8398	6703	7550
1987	191.77	918	0.0570	3362	16095	7767	6841	7304
1988	326.86	2656	0.1221	2676	21744	6181	9241	7711
1989	466.88	3410	0.1466	3185	23263	7357	9887	8622
1990	543.62	3989	0.1857	2928	21486	6764	9132	7948
1991	675.47	4290	0.2277	2966	18838	6852	8006	7429
1992	736.88	4331	0.2630	2801	16465	6471	6998	6734
1993	742.08	4810	0.2887	2570	16661	5938	7081	6509
1994	626.48	5160	0.3088	2029	16710	4686	7102	5894
1995	861.59	5670	0.4169	2067	13601	4774	5780	5277
1996	1362.63	9800	0.5602	2432	17494	5619	7435	6527
1997	1271.24	11600	0.6757	1881	17166	4346	7296	5821
1998	1470.79	11690	0.7834	1877	14922	4337	6342	5340
1999	1530.68	12230	0.9133	1676	13391	3871	5691	4781
2000	1505.94	12270	1.0000	1506	12270	3479	5215	4347
2001	1612.09	13110	1.0637	1516	12325	3501	5238	4370
2002	1785.12	11600	1.1172	1598	10383	3691	4413	4052
2003	1822.15	13030	1.1680	1560	11156	3604	4741	4173
2004	1897.44	14560	1.2227	1552	11908	3585	5061	4323
2005	1910.08	15930	1.2715	1502	12528	3470	5325	4397
2006	1967.03	16590	1.3177	1493	12591	3448	5351	4400
2007	2774.72	16690	1.3699	2025	12183	4679	5178	4928
2008	3038.12	18060	1.4401	2110	12540	4873	5330	5101
2009	3166.04	18600	1.5164	2088	12266	4823	5213	5018
2010	3399.52	19180	1.5844	2146	12106	4956	5145	5051
2011	3917.19	18840	1.6333	2398	11535	5540	4902	5221

Source: Author calculations based on data from SAGARPA (2012).

Chapter 3
A Mixed Multinomial - Conditional model of land use decisions.

3.1 Model description.

Since the work of Chomitz and Gray (1996), spatially explicit discrete choice models of land use decisions have been one of the most commonly used approaches in the study of landscape dynamics. The typical formulation considers that agents' land allocations are mainly driven by the expected payoffs of the land use choices available in a particular geographical area, and that such payoffs are affected by parcel specific variables such as slope, distance to nearest markets, soil quality, etc., as well as by socioeconomic, technological, cultural and ecological factors (Geist & Lambin, 2002; Redman et al., 2004). The state of those land use drivers at a particular time $t = 1,...,T$, for $T \leq \infty$, determines the choice set, C_{it}, of land uses available to decision maker i at different t periods. For a specific region composed of a set of agents η and a finite set of time periods τ, the analyst can define a finite choice set $J = \bigcup_{i \in \eta}^{t \in \tau} C_{it}$ that includes all the feasible land uses in the region.

Generally, it is assumed that at time *t*, an economically rational agent *i* chooses the land use with the highest expected payoff both at time *t* and in future periods. An additional common assumption in the land use literature is that at every period t, agent i

knows the value of the per period payoff, u_{ijt} $\forall j \in J$, but the researcher can only observe the agent's land use choice d_{ijt} (with $d_{ijt} = 1$ if land use *j* is selected, and 0 otherwise), and the state of some of the variables that determine the payoff $\forall j \in J$. In the context of random utility theory u_{ijt} can be decomposed into an observable *systematic component* (or *representative utility*), S_{ijt}, and a random term, ε_{ijt}, that accounts for the effects of factors that are unobservable for the analyst that impact the payoff levels for each land use in the choice set (Rust, 1994).

Typically the systematic component of the utility is modeled as a profit function that depends on input and output prices, expected yield, and parcel characteristics that impact the productivity or quality of the uses attached to each parcel (Baerenklau et al., 2012; Chomitz & Gray, 1996). Following Baerenklau et al. (2012) we can consider that the net profit from land use j at time *t* can be represented as $v_{jt} = p_{jt} y_{jt} - \mathbf{q}_{jt}{}'\mathbf{c}_{jt}$ where the revenue from land use j is computed using information on the price per unit of output *j* at time *t*, p_{jt}, and the yield per parcel in that period , y_{jt}; the production costs are computed as the dot product of a vector of input prices, $\mathbf{q}_{jt}$, and a vector of quantities of each input required by land use *j*, $\mathbf{c}_{jt}$. A vector of parcel characteristics $\mathbf{k}_i$ can be incorporated in the analysis to estimate its impact on the payoffs for each land use. If we define u_{ijt} as the random variable: $u_{ijt} \equiv V_{ijt} = S_{ijt}\left(v_{jt}, \mathbf{k}_i\right) + \varepsilon_{ijt}$, and represent the portion of the payoff that can be observed by the analyst as $S_{ijt}\left(v_{jt}, \mathbf{k}_i\right) = \alpha_j + \beta'_j v_{jt} + \boldsymbol{\omega}'_j \mathbf{k}_i$ with α_j, β_j and $\boldsymbol{\omega}_j$ representing alternative specific coefficients that can be used to estimate the marginal effects of changes in prices and parcel characteristics in the landscape configuration. We can then represent agent's *i* payoff from choosing land use *j* at any *t* period as $V_{ijt} = \alpha_j + \beta'_j v_{jt} + \boldsymbol{\omega}'_j \mathbf{k}_i + \varepsilon_{ijt}$.

As stated in the previous paragraphs, economically rational agents are assumed to select land use *j* if, and only if, its payoff is higher than the payoff from any other alternative land use available in their choice set. Dropping the time index to simplify the notation this can be stated as $d_{ij} = 1$ iff $V_{ij} > V_{il} \ \forall j, l \ \in J \wedge j \neq l$. Following Croissant (2008) we can analyze the implication of this utility comparison to see the effects of setting alternative specific coefficients for the alternative specific variables (e.g., output prices), and parcel specific coefficients for the variables that represent agent heterogeneity (e.g., distance from a parcel to urban areas, soil quality, landowner's education). The following derivation provides some insights of the implication of the assumed utility comparison process followed by decision makers. If land use *j* is preferred over land use *l* by agent *i*, this is an indication that,

$$
\begin{aligned}
V_{ij} > V_{il} \Rightarrow \alpha_j + \beta_j' v_j + \boldsymbol{\omega}'_j \mathbf{k}_i + \varepsilon_{ij} &> \alpha_l + \beta_l' v_l + \boldsymbol{\omega}'_k \mathbf{k}_i + \varepsilon_{il} \\
\alpha_j - \alpha_l + \beta_j' v_j - \beta_l' v_l + \boldsymbol{\omega}'_j \mathbf{k}_i - \boldsymbol{\omega}'_k \mathbf{k}_i &> \varepsilon_{il} - \varepsilon_{ij} \\
\alpha_j - \alpha_l + \beta_j' v_j - \beta_l' v_l + \left(\boldsymbol{\omega}_j - \boldsymbol{\omega}_k\right)' \mathbf{k}_i &> \varepsilon_{il} - \varepsilon_{ij}
\end{aligned}
$$

This expression shows that the coefficients for the parcel specific variables $\mathbf{k}_i$ and the constant terms need to be alternative specific and that only their differences matter. This implies that we can estimate coefficients on parcel characteristics for all but one of the land uses contained in the choice set. The common normalization criterion is to set one of the coefficients equal to zero. On the other hand, we can observe that coefficients for all the alternative specific variables can be computed without requiring normalization.

Assuming that the random terms are independent Gumbel distributed variates, the probability of agent *i* selecting the observed land use in the dataset can be computed as,

$$\Pr\left(d_{ij}=1\middle|v_j,\mathbf{k}_i;\alpha_j,\beta_j,\boldsymbol{\omega}_j\right)=\frac{e^{S_{ij}\left(v_j,\mathbf{k}_i\right)}}{\sum_{l\in J}e^{S_{il}\left(v_l,\mathbf{k}_i\right)}}$$

Assuming that the sample data is composed by independent land use decisions, the probability of each agent in the sample selecting the observed land use can be represented by the likelihood function $\mathrm{L}(\varphi|\mathrm{v}_j,\mathbf{k}_i,\mathrm{d}_{ij})=\prod_{i=1}^{N}\prod_{j}\Pr_{ij}^{d_{ij}}$ where $\varphi=\beta_j\bigcup\boldsymbol{\omega}_j\bigcup\alpha_j\ \forall\ j\in J$.

Since the product of the individual probabilities over all the land options can generate a small number that can be affected by rounding errors it is more convenient to use the log of the likelihood function to compute the parameter estimates (Enders, 2010). Following Train (2009), the log-likelihood is,

$$\begin{aligned}
\mathrm{LL}(\varphi|\mathrm{v}_j,\mathbf{k}_i)&=\sum_{i=1}^{N}\sum_{j\in J}\mathrm{d}_{ij}\ln\Pr_{ij}\\
&=\sum_{i=1}^{N}\sum_{j\in J}\mathrm{d}_{ij}\ln\left(\frac{e^{\alpha_j+\beta_j^{'}v_j+\boldsymbol{\omega}_j^{'}\mathbf{k}_i}}{\sum_{l\in J}e^{\alpha_l+\beta_l^{'}v_l+\boldsymbol{\omega}_l^{'}\mathbf{k}_i}}\right)\\
&=\sum_{i=1}^{N}\sum_{j\in J}\mathrm{d}_{ij}\left(\alpha_j+\beta_j^{'}v_j+\boldsymbol{\omega}_j^{'}\mathbf{k}_i\right)-\sum_{i=1}^{N}\sum_{j\in J}\mathrm{d}_{ij}\ln\sum_{l\in J}e^{\alpha_l+\beta_l^{'}v_l+\boldsymbol{\omega}_l^{'}\mathbf{k}_i}
\end{aligned}$$

The marginal effect of a change in the revenue of land use *j* on the probability of selecting that particular land use can be represented as $\frac{\partial \Pr_{ij}}{\partial v_j} = \beta_j \Pr_{ij}\left(1-\Pr_{ij}\right)$. The marginal effects of changes in parcel specific variables can be estimated using the formula $\frac{\partial \Pr_{ij}}{\partial k_i} = \Pr_{ij}\left(\omega_j - \sum_{l\in J}\omega_l \Pr_{il}\right)$.

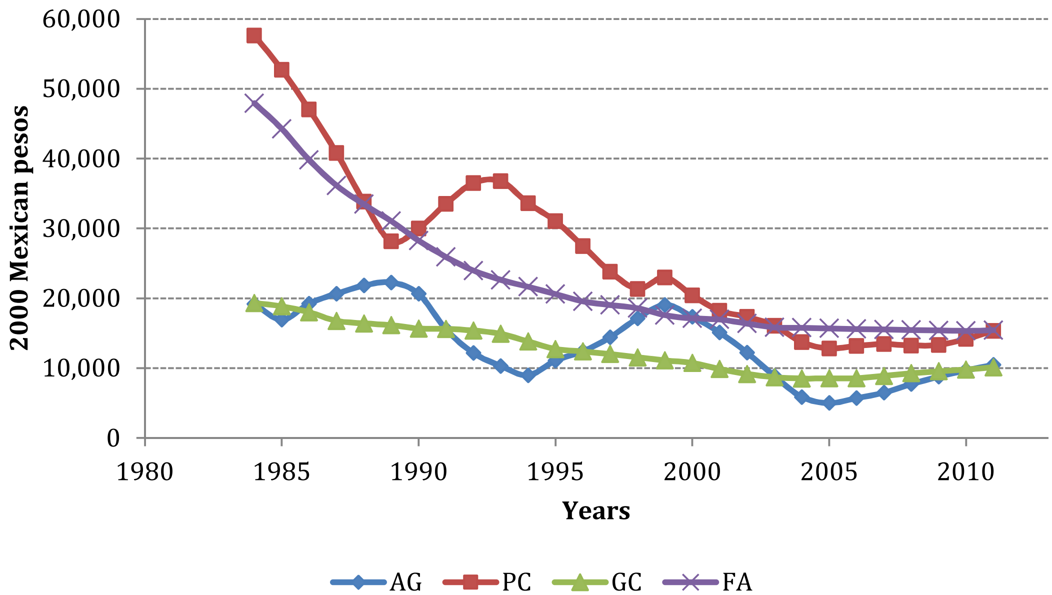
3.2 Empirical application.

The described modeling framework was used in previous analysis implemented by Ellis et al. (2010) and Baerenklau et al. (2012) to identify and analyze land use and land cover change trends in two Mexican coffee growing municipalities during the period 1984-2006. Those authors use remotely sensed data, coupled with geographic and socioeconomic information to implement a multinomial logit analysis of the marginal effects of spatial and economic factors that affect land use choices between Agroforests (AG), Perennial crops (PC), Grass and corn (GC), and Fallow lands (FA) in the study area, and to test the welfare effects and landscape configuration resulting from simulated policies that increase the expected revenue of agroforestry parcels. Those studies, that constitute an example of a standard implementation of a spatially explicit model of land use decisions, are considered as a starting point in the mixed multinomial-conditional logit (MMCL) analysis implemented in this chapter. The MMCL modeling approach differs from the multinomial logit analysis implemented by Ellis et al. (2010) and Baerenklau et al. (2012) in one key assumption, it considers that the portion of the expected land use payoff that is determined with market information, should be modeled as an alternative specific variable, especially if that component is computed using market prices and average yield values that are independent of parcel characteristics. On the other hand, if there is not market information on output prices, and land use revenues are considered to be endogenously determined by parcel characteristics (see Chomitz and Gray (1996), Blackman et al. (2008), Nelson et al. (2004) and De Pinto and Nelson (2008) for some examples); one of the coefficients of the variables that capture the effects of prices on agents' decisions needs to be normalized for model identification and in those cases a multinomial logit model is an appropriate modeling mechanism.

In the study region, farmers on average receive similar prices for their yields independently of the location of their parcels. Obviously, distance to markets and yield quality directly affect the net revenue received by landowners but those factors are assumed to be controlled in the MMCL by the parcel specific characteristics included in the dataset, and the unobservable components. An adequate computation of the alternative specific component of the payoff for land use *j* requires knowledge about the age of the current land use and about the price expectation process followed by the decision maker. Even though the sample design allows approximating the age of the land uses in each parcel, that information is not incorporated in the MMCL analysis since the purpose of this chapter is to implement a standard economic model of discrete land use decisions to highlight the advantages and limitations of this approach and accounting for age of the land uses is not commonly found in similar spatially explicit models. With regard to the price expectation process, farmers in the study area appear to be irresponsive to short-term fluctuations in land use profitability, especially when land use change requires at significant up-front investment to switch to a different use.

Nevertheless, if the price trends continue during several years agents become more likely to replace the use of their land for a more profitable use. I follow Baerenklau (2012), to consider that landowners not only use the price information available in the current period but also the price signals observed during the previous past five years to estimate expected revenues. I construct a moving average time series revenue data with inflation-adjusted price indexes, and an average expected yield for each relevant j land use during 25 years of production, using the formula, $p_{jt}y_{jt} \equiv \frac{1}{5}\sum_{s=t-4}^{t} p_{js}\overline{y}_j$ with $\overline{y}_j = \frac{1}{25}\sum_{i=1}^{25} y_j$ for land use j corresponding to a land use that is different to the observed at the beginning of the decision period, and $\overline{y}_j$ equal to the maximum yield for a mature plantation to estimate the expected revenue of continuing the same land use during an additional year. The limitations of this procedure are obvious, but commonly implemented in the context of myopic discrete choice static models. A more structured approach is implemented in chapter five using a dynamic model of land use decisions that consider that yield is dependent on the estimated age of each land use at the parcel level. Figure 3.1 shows the trends observed in the expected revenue from mature land uses in a 2-hectare parcel, which is the parcel size that we consider in the econometric analysis for two reasons: it is a farm-size that can be managed by one person giving a meaningful value to the FA price index, and because most of the shade grown coffee plantations are around that parcel size (Escamilla Prado, 2007).

Figure 3.1. Expected revenue from mature land uses in a 2-hectare parcel.



On the other hand, empirical observations in the study region indicate that the agricultural production technology has been labor intensive and without significant improvement during the last decades. Those observations appear to be supported by studies developed in similar regions. For instance, Gay et al. (2006) in a study of coffee production in the state of Veracruz, Mexico, finds that labor cost can account for up to 80% of the total production costs. Similarly, Albers et al. indicate that (2006) shade-grown coffee plantations require few purchased inputs such as pesticides and fertilizers. Therefore, we consider that production costs for each land use j, $\mathbf{q}_{jt}'\mathbf{c}_{jt}$, are relatively constant during the period of analysis and captured by the constant term and parcel characteristics in the econometric model.

Summarizing, the behavioral decision-making process modeled with the MMCL approach assumes that agents:

a) use current price information to update a five-year moving average price expectation process to estimate future price realizations.
b) use an average of the expected yearly output per land use during the first 25 years of a new land use to estimate the expected yield for different land uses to the observed at the beginning of the decision period.
c) utilize the expected yield for a mature land use as the expected yield if the current land use is not changed during the decision period.

Additionally, it is assumed that average production costs and switching costs are relatively constant during the period of analysis and its effects captured in the constants, unobservable variables and parcel specific information.

3.3 Results and discussion

3.3.1 Parameter estimates and marginal effects.

The mixed multinomial – conditional logit model of land use choices was estimated using the mlogit R package (Croissant, 2008). The estimated coefficients of the variables considered in the analysis are shown in table 3.1. We can observe that all the parameter estimates of the revenue variable are statistically significant at the 10% level. Given that for alternative specific variables, signs in the parameter estimates can be interpreted as directions of marginal effects, we can observe that the econometric model produces counterintuitive parameter estimates for the revenue associated with perennial crops, grass and corn, and the fallow category. The agroforestry revenue coefficient has the appropriate direction although its significance level is the lowest among the alternative specific coefficients.

Table 3.1 Mixed Multinomial – Conditional Logit parameter estimates.

	Agroforestry		Perennial Crops		Grass and Corn		Fallow	
Revenue	**0.02689**	.	**-0.05638**	*	**-0.09415**	***	**-0.13432**	***
	0.01462		0.02253		0.02200		0.01358	
Slope	**0.41628**	***	**-0.21461**	*	**-0.14862**			
	0.09032		0.09896		0.14152			
Distance to market	**0.65282**	***	**0.92505**	***	**1.03446**	***		
	0.18172		0.18245		0.21731			
Distance to road	**1.61612**	***	**1.43768**	***	**1.20529**	**		
	0.34001		0.33020		0.39877			
Poverty	**-0.05792**		**-0.31780**		**0.67688**	*		
	0.25020		0.24742		0.32355			
Soil texture	**0.04850**		**-0.57309**	**	**-0.37084**			
	0.17843		0.17651		0.25894			
Elevation	**5.17566**	***	**-1.37557**	.	**-2.28093**	*		
	0.77509		0.76828		1.14876			
Population	**-0.28962**	***	**-0.55676**	***	**-0.58732**	***		
	0.07456		0.07980		0.13211			
Constant	**-5.49572**	***	**-0.18972**		**-1.48418**	.		
	0.62667		0.58782		0.85720			

Notes: The parcel specific coefficients of the fallow category were normalized to zero for model identification. The corresponding standard errors are shown below each parameter estimate. Significance codes: '***' significant at the 0.1% level; '**'significant at the 1% level; '*' significant at the 5% level; '.' Significant at the 10%.

On the other hand, since the signs of the estimated coefficients of the parcel specific variables do not indicate the direction of their marginal effects, I estimate how changes in each of the variables used in the analysis affect the land use probabilities for each parcel and then averaged the results across all observations, the mean marginal effects, standard deviation and expected signs as described in chapter 2 are shown in table 3.2. In general, the directions of the marginal effects are consistent with the multinomial logit analysis of land use decisions in this region developed by Ellis et al. (2010) and Baerenklau et al. (2012). The results indicate that parcels at higher elevations have greater probability of being selected for agroforestry production. This is consistent with agroecological/productivity requirements for coffee production, which constitutes the main component of the agroforestry land use category. The negative and statistically significant effect of elevation on the probability of selecting perennial crops is also consistent with requirements to produce the crops included in that land use category. Slope has a positive and statistically significant effect on the probability of a parcel being devoted to agroforestry production. This is consistent with informal observations in the study region that indicate that parcels with relatively high slope are mostly devoted to coffee production in areas very difficult to cultivate. The results also indicate an inverse and statistically significant effect of the distance from a parcel to the nearest market in the probability of a parcel being selected for agroforestry or perennial crop production. The marginal effects of distance to the nearest market in the probability of selecting grass and corn or fallow are coherent with the intuition that if a parcel is far from a market it would be better to let it fallow or to grow subsistence crops. The marginal effects that correspond to the poverty index have the expected signs indicating than cash crops are more likely to be selected in areas with lower poverty levels, and that subsistence crops would be preferred in poorer areas.

Table 3.2. Marginal effects.

		Expected sign	Estimate	Standard Deviation
Revenue	AG	+	0.0049	0.0016
	PC	+	-0.0055	0.0048
	GC	+	-0.0182	0.0055
	FA	+	-0.0086	0.0065
Slope	AG	+	0.1019	0.0347
	PC	-	-0.0329	0.0328
	GC	-	-0.0674	0.0345
	FA	+ -	-0.0016	0.0077
Distance to market	AG	-	-0.0425	0.0198
	PC	-	0.0161	0.0162
	GC	+	0.0838	0.0404
	FA	+	-0.0574	0.0460
Distance to road	AG	-	0.0891	0.0384
	PC	-	0.0161	0.0266
	GC	+	-0.0171	0.0506
	FA	+	-0.0881	0.0635
Poverty index	AG	-	-0.0746	0.0503
	PC	-	-0.0585	0.0558
	GC	+	0.1517	0.0473
	FA	+	-0.0186	0.0241
Soil texture	AG	+	0.0770	0.0289
	PC	+	-0.0419	0.0378
	GC	-	-0.0520	0.0324
	FA	+-	0.0170	0.0161
Elevation	AG	+	1.2598	0.4277
	PC	-	-0.2633	0.3147
	GC	+-	-0.9786	0.4083
	FA	+-	-0.0179	0.0979
Population	AG	+	0.0402	0.0158
	PC	+	-0.0158	0.0152
	GC	-	-0.0556	0.0254
	FA	-	0.0312	0.0256

Expected direction codes: ‘ + ‘ indicates that a positive marginal effect is expected, ‘ – ‘ indicates that a negative marginal effect is expected, ‘ + - ’ indicates that the marginal effects can go in either direction. See chapter 2 for a detailed description of the expected direction of the marginal effects.

3.3.2 Predictive power.

The estimated probabilities of agent *i* selecting land use $j \in J$ were used to detect the land use with the highest probability of being selected. On average the model accurately predicts land use choices in 59% of the parcels which is relatively close to the 52% predictive accuracy with the multinomial logit model used by Baerenklau et al. (2012) in the same region. Figure 3.2 shows the prediction accuracy of the estimated land use decisions during the years for which we have land use observations. We can observe that the model does a relatively good job predicting land use decisions during the first two periods in the dataset, after that the accuracy falls below the mean. Figure 3.3 shows that the model does not predict fallowing in any period of the analysis, it overestimates agroforestry and grass and corn choices, and it underestimates the proportion of land devoted to perennial crops.

Figure 3.2. Parcel-level accuracy of land use predictions

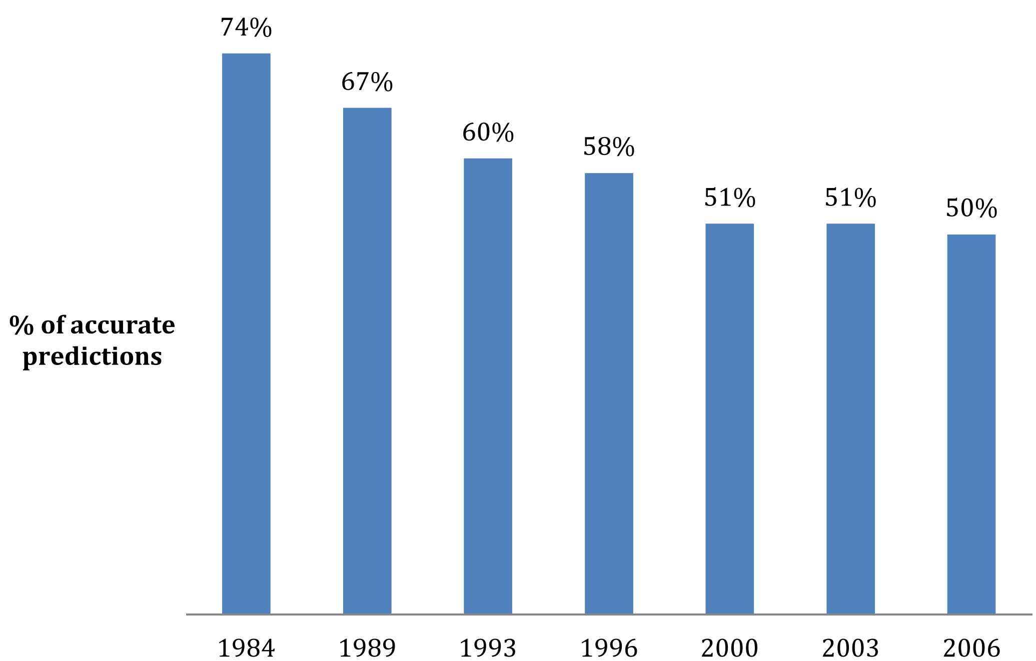
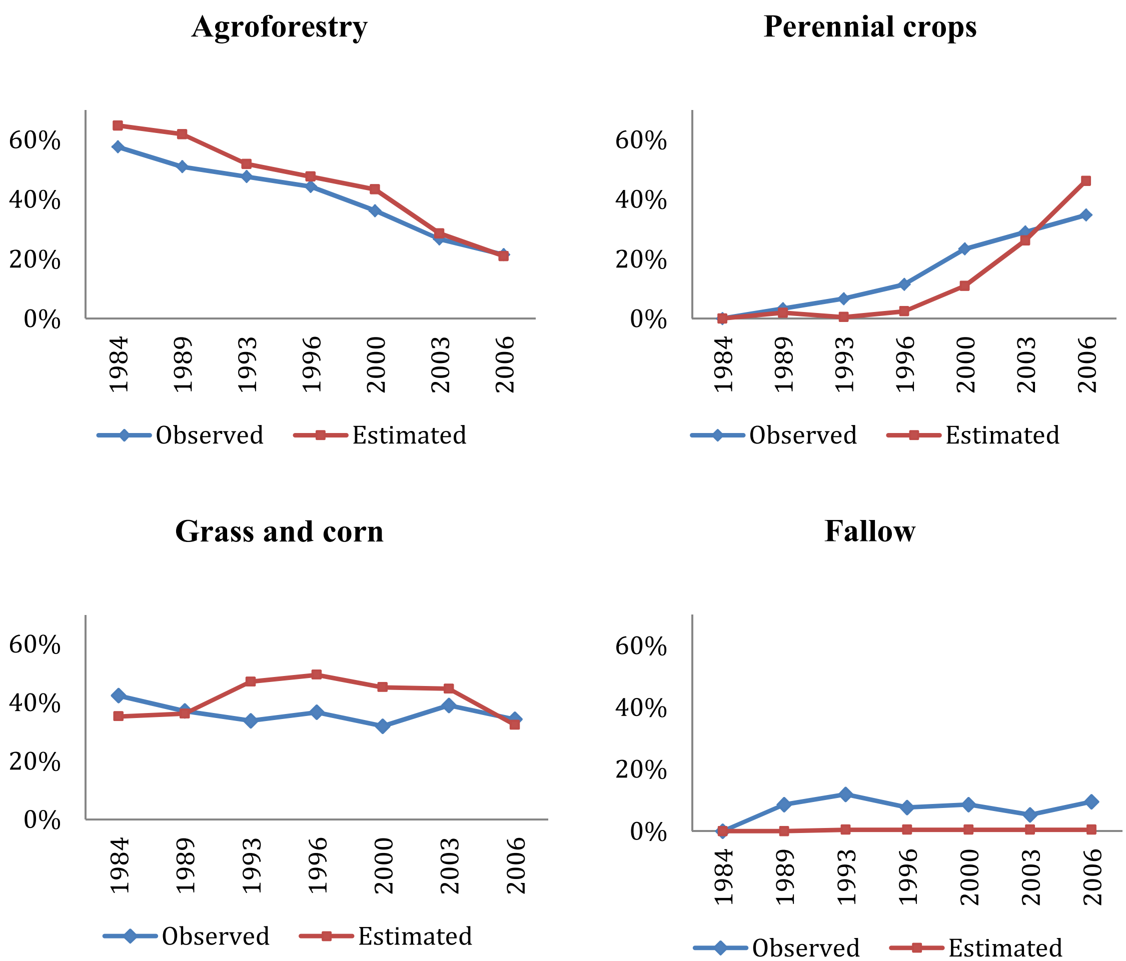
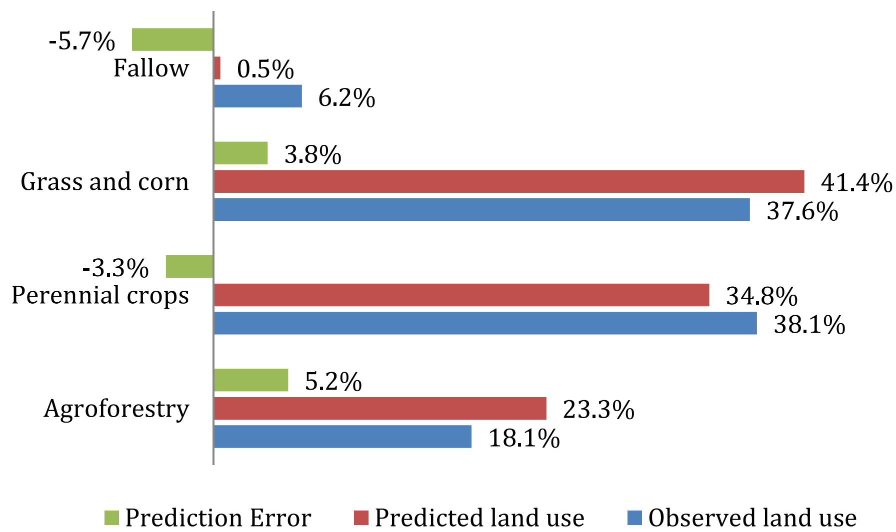


Figure 3.3. Percentage of observed and estimated land uses (1984 – 2006).



To evaluate the out-of sample predictive power of the model, we use the parameter estimates and the corresponding values of the explanatory variables in 2011 to estimate land use choices in that year. To test the accuracy of the predictions we use Google Earth to identify the land uses for the sample parcels in 2011. The declining trend in the prediction accuracy of the model observed in figure 3.2 continues through 2011 since the model accurately predicts only 37.6% of the observed land use decisions during that year. Nevertheless, if we analyze the predicted land use proportions, the model generally does a decent job of predicting the land distribution between the land use categories (Figure 3.4).

Figure 3.4. Percentage of observed and estimated land uses in 2011.



3.4 Summary

In general terms the MMCL model produces marginal effect estimates that are consistent with the expected directions described for the parcel specific variables that account for population, elevation, soil texture, and slope. Nevertheless, this reduced form static model does not produce theoretically consistent parameter estimates for the revenue variable associated with three of the land uses. The model accurately captures the impact of declining revenue on the percentage of agroforestry parcels but it cannot provide an economically rational explanation for the increasing proportion of perennial crops and the concomitant decrease in the associated revenue. The predictive performance of the model is within the range of similar studies. Although the per-parcel prediction accuracy of the model decreases through time, potentially indicating that this model may be a good mechanism for short term predictions but perhaps not useful for extended time periods.

Unfortunately, discrete choice static models have inherent limitations that complicate a realistic approach to the behavioral process followed by landowners. One of the key assumptions of myopic discrete choice models is that decision makers only use historical data and current information to make their choices. In reality, landowners not only consider the current state of the world but also incorporate forward-looking behavior and some type of risk analysis to evaluate the expected returns and consequences derived from their decisions. Furthermore it has been shown that the study of inherently dynamic economic problems within static frameworks can generate biased and theoretically inconsistent estimates and consequently less accurate policy recommendations (Baerenklau & Provencher, 2005; Hicks & Schnier, 2006). Therefore, the development and implementation of agent-based models that account for the dynamic nature of land use decisions is necessary both to evaluate the magnitude of the bias in estimates derived from commonly used static models and to develop more accurate policy recommendations.

Additionally, in the tested model it is possible that landowners use an expectation-formation process to estimate future prices that is different than the simple moving average approach used in this chapter which may produce inaccurate parameter estimates for the revenue variable. It is also possible that some of the parcels classified in the agroforestry category are in fact fallowed lands and that some of the fallowed lands are part of rotational constraints and have not been abandoned, but since land use classification methods cannot identify such situations the parameter estimates and marginal effects may be biased.

Despite all the aforementioned drawbacks, which are common to this type of modeling approach, discrete choice static models continue to be used to analyze agent based land use decisions. The next chapters present alternatives to some of the limiting assumptions or problems faced by the type of model implemented in this chapter.

Chapter 4
Discrete choice analysis of land use decisions with misclassified data.

4.1 Introduction

Advances in remote sensing technologies during the last few decades have allowed the characterization of landscape dynamics at different temporal and spatial scales. Particularly, the increasing availability of satellite imagery has contributed to the generation of global, regional and local land cover and land use information. Typically, the mapping of remotely sensed data is implemented using supervised or unsupervised classification methods, a combination of both approaches, or through object-based classification. Data collected at specific geographic locations is often used to calibrate the classification algorithms and to reduce classification errors. Unfortunately, several conditions might generate misclassifications. In some cases, the resolution or quality of the remotely sensed data complicates the classification process. For instance when the image has a high percentage of its area covered with clouds or when the pixel size is very large that only a coarse land use classification can be implemented. In other cases, the complexity of the landscape configuration limits the accuracy of the pixel-based classification algorithms used to categorize land uses with similar spectral signatures (e.g., land uses with similar biomass density). The quest for better procedures to extract land use information from remotely sensed data has motivated a significant amount of research during the last decades to improve or develop new classification algorithms.

Nevertheless, after reviewing more than five hundred classifications of satellite images published during the period 1989-2003, Wilkinson (2005) reports that such efforts have not materialized into a significant improvement in the accuracy of pixel-based land use classifications. More recently, the increasing availability of high resolution spatial data and computational advances have facilitated the development of more sophisticated classifiers that not only use pixel information but also pattern recognition algorithms to improve the classification of remotely sensed data although their effectiveness has not been extensively tested (see Du et al. 2012; Hongzan et al. 2011 for some examples).

I start with the assumption that classification errors are an intrinsic component of the land use classification process but that statistical methods can be implemented to improve the accuracy of land use datasets. My aim is to investigate the impact of inaccurate land use data on the analysis of land use decisions and derived policy recommendations. This is a particularly important issue in the dataset of land use information described in Chapter 2. The categorization of land uses in the coffee growing region of Atzalan, Veracruz, Mexico proved to be a complex task. The vegetation density of forested areas, agroforestry parcels, or land that has been fallowed during several years produces similar spectral values that can mislead even the most sophisticated pixel-based classifiers. In our dataset it was practically impossible to accurately classify in separate categories those land uses using only remotely sensed data and ground truthing.

To address this issue, I implement a post-classification strategy that simultaneously detects misclassified land use decisions using the information generated from the remotely sensed data, and incorporates those corrections into a latent multinomial logit (LMNL) land use model. The following section presents a review of relevant research that constitutes the basis of the modeling approach implemented in this chapter. The third section presents a methodological description of the LMNL model implemented in the context of agent-based land use decisions. The fourth section presents the results of the empirical application, including a comparison of the parameter estimates and marginal effects derived from a standard multinomial logit model implemented with the original dataset, with those derived from the LMNL model that corrects for misclassified. Finally, the last section presents the conclusions as well as a discussion of the limitations of the analysis.

4.2 Literature Review

Since the seminal work of Dempster et al. (1977), the expectation maximization (EM) algorithm has been used to generate parameter estimates in probabilistic models with incomplete or misclassified data. This is typically done by associating an incomplete data problem with a complete-data problem for which maximum likelihood estimation is manageable (McLachlan & Krishnan, 1997). An iterative process between the expectation step (E-step) and the maximization step (M-step) is the basis of the EM algorithm. The E-step computes the expectation of the missing data conditional on the given set of incomplete information and initial values of the parameters to be estimated. The M-step uses those conditional expectations in the place of the missing information to "complete" the dataset and estimate the parameters that maximize the likelihood function for the "complete-data" problem. The parameter estimates produced in the M-step are used as updated initial values of the coefficients in the E-step and the process is repeated until the likelihood converges to a local maximum (McLachlan & Krishnan, 1997; Zhai, 2007).

Given its extensive implementation across several disciplines the EM procedure is considered to be a highly desirable algorithm in data mining (Wu et al., 2007). In the context of land use and land cover mapping the EM algorithm has been used to refine unsupervised classification methods (Chardin & Perez, 1999; Yang, Peng, Xia, & Zhang, 2013); to estimate the pixel values of portions of remotely sensed imagery that are missing due to the presence of clouds during the time of data collection (Melgani, 2006); and to improve the classification accuracy of pixels that include information of more than one land use category (Susaki, J., & Shibasaki, 2000). Nevertheless, to our knowledge the EM algorithm has not been used to analyze the impact of misclassified data on agent based land use models, a task that can be accomplished using a latent multinomial logit model (LMNL).

The LMNL model uses a nesting structure to classify observed choices as branches that may or may not contain misclassified observations. In other words, a discrete choice dataset can be decomposed into $1,\ldots, n$ branches representing the choice set available to the decision maker. Each k branch in the choice set contains a stem k that groups data that is accurately classified and may contain up to $n - 1$ stems that cluster misclassified observations that should be included in a different branch. Caudill (2006), describes the methodology that can be used to produce parameter estimates with a dataset containing misclassified dependent variables, as is the case here. The procedure is based on a transformation of the standard multinomial logit likelihood function into a missing data formulation to which the EM algorithm can be applied. Caudill and Mixon Jr. (2005) use this methodology to study misleading responses provided by students in a survey that collects information on undergraduate cheating behavior. Caudill et al. (2005) use the LMNL approach to estimate the proportion of fraudulent claims for car damage that are inaccurately classified as honest by a Spanish insurance company (Caudill et al., 2005). The LMNL methodology is also used by Caudill (2006) to estimate the impact of misclassified observations on an analysis of hidden unemployment in six European countries. The author finds statistical evidence to argue that workers who are wrongly reported as own-account self-employed in the survey represent an additional 0.5% of hidden unemployment in the studied economies. More recently, the study by Caudill et al. (2011) uses an unconstrained version of the LMNL model to analyze hypothetical bias (the situation in which stated willingness to pay is higher than the actual willingness to pay) in a contingent valuation problem. The LMNL methodology offers a straightforward procedure to handle misclassified land use information. The specifics of the particular model implemented in this chapter are presented in the following section.

4.3 Empirical Problem and Methodology

Most of the spatially explicit models of land use decisions in rural areas have been focused on analyzing how driving forces of deforestation reconfigure pristine landscapes affecting the provision of environmental services (Andersen, 1996; Chomitz & Gray, 1996; Geist & Lambin, 2002; Puri, 2006). Nevertheless, the growing recognition that agroforestry production systems can provide forest-like services as well as biodiversity corridors between patches of forested or protected areas have highlighted the need of understanding land use decisions in those areas (Ávalos-Sartorio & Blackman, 2010; Bhagwat et al., 2008; Dinata Putra et al., 2005; Huang et al., 2002; Kursten, 2000; Schroth, 2004; Shanker & Solanki, 2000; Swallow et al., 2006). Worldwide, shade-grown coffee plantations are one of the most important agroforestry production systems not only for their socio-economic relevance providing livelihood opportunities to many farmers (Albers et al., 2006; Aoki & Suvedi, 2012; Blackman et al., 2012; Jordan-Garcia, Collazo, Borkhataria, & Groom, 2012; Oxfam, 2002), but also for the ecological services that those plantations provide (Escamilla Prado, 2007; Messer, Kotchen, & Moore, 2000). In Mexico small-scale farmers across the country depend upon shade grown crops, with coffee being the leader both in terms of cultivated land area and value of production. Escamilla-Prado (2007) reports that around 3 million people in Mexico depend on coffee-related activities and that approximately 90% of the coffee-cultivated area lays under diversified shade. Unfortunately, the steady decline in the international coffee prices observed during the 1990's and first years of the 2000's forced coffee farmers to find alternative sources of income. Some farmers opted for coffee certification schemes to obtain a price premium for implementing environmentally friendly production techniques, while others decided to clear their coffee plantations to transition to a different land use or abandoned their plantations to look for employment opportunities in other economic sectors and/or geographical locations (Blackman et al., 2008; Lewis & Runsten, 2008; Nava-Tablada & Martínez-Camarillo, 2012).

As stated in chapter 2, the spatially explicit dataset of land use decisions used in this chapter was constructed using maximum likelihood supervised classification with ground thruthing on six Landsat images for the years 1984, 1989, 1993, 1996, 2000, 2003, and one Spot image for the year 2006. The spectral information contained in the remotely sensed data was used to map the landscape in the geographical region under analysis into six land use categories: secondary forests, shade grown coffee, banana, citrus, pasture, and corn. To reduce classification errors disaggregated land uses were integrated into three general land use categories: agroforestry (AG) which is composed by shade grown coffee plantations and secondary forest; perennial crops (PC) integrated by citrus or banana plantations; and grasslands and cornfields (GC). The main criteria used to construct the aggregated land use categories are that their components share similar tree canopy density, profitability and conversion costs. Additionally, understanding that under some circumstances agents would prefer to let their land fallow and look for alternative income sources, for instance when the land use profitability falls below some threshold level, I constructed an additional land use category composed by fallowed parcels (FA). This land use type was assigned to some parcels after analyzing the sequence of land use decisions produced with the remotely sensed data. I consider that a temporary land use transition that last at most six years (roughly two observation intervals) from GC or PC to AG and then back to the previously observed land use indicates that that parcel was in fact fallowed during the period detected as AG. An example may clarify the procedure. Consider that after classifying the remotely sensed data the land use in parcel *s* is classified as GC during 1996, AG during 2000 and again GC in 2003. This land use sequence is not logical either by economic or biological reasoning. In cases like this I consider that parcel *s* was in fact fallowed during 2000 and that the classifier algorithm categorized the land use as AG after detecting an increase in biomass that was likely generated because the landowner forwent maintenance activities in that parcel. This procedure allowed us to identify potential fallowed parcels mainly in the GC category since temporary land use transitions between PC and AG represent less than 0.15% of the sample dataset. This is a relevant drawback of the procedure to detect FA parcels in the PC category since land abandonment of citrus or banana plantations is not common given the potential impacts on future yields that the lack of maintenance activities may have.

Nevertheless, there are some undeniable complications in the method that I use to construct the FA category. On the one hand, the procedure makes it practically impossible to detect AG parcels that are in fact fallowed plots during any period. This is potentially a relevant issue, since Albers et al. (2006) reports that at least 75% of farmers in a coffee growing region in Oaxaca, Mexico forwent maintenance activities during the coffee crisis period (1990 – 2004). Empirical observations of coffee growing parcels in the study region indicate that abandoned AG parcels were common at least after 2000. On the other hand, the transition observed in some parcels between GC and FA may be part of a rotational production system used to recover soil productivity (Adiku, Kumaga, Tonyigah, & Jones, 2009; Kolawole, Salako, Idinoba, Kang, & Tian, 2005; Tian, Salako, Kolawole, & Kang, 1999). This means that it is possible that some of the parcels classified as FA are in fact GC following a rotational scheme and that the land use of those parcels has not actually changed. Alternatively, it is also possible that grasslands or cornfields with a relative increase in biomass are in fact parcels that have not received maintenance activities during the period in which the remotely sensed data was collected. Unfortunately, these types of misclassification problems cannot be addressed using algorithms based on spectral information. Nevertheless, we can use the LMNL model to estimate the probability that an AG parcel is actually fallowed as well as the probability that a parcel classified as FA is in fact a GC plot that has been temporarily abandoned.

The approach used to detect misclassified land use decisions is framed in the context of the widely implemented random utility discrete choice models. As described in Chapter 3, these types of models consider that variations in the socioeconomic, cultural and ecological systems drive land use changes through their impacts on the expected payoffs that landowners use to determine their decisions. Let $\mathbf{X}_i$ represent the matrix of observable variables that determine the expected net revenue for all the optional land uses in the choice set J for agent *i* with $i = 1, ..., n$; β_j represent a vector of coefficients for the explanatory variables that affect the payoff of land use *j*; and α_j represent the alternative *j* specific constant term; under the assumption that the unobservable components that determine land use *j* payoffs are independent Gumbel distributed variates the probability of agent *i* selecting land use *j,* can be computed as

$$\Pr_{ij}\left(d_{ij} = 1 \middle| \mathbf{X}_i, \beta_j, \alpha_j\right) = \frac{e^{\alpha_j + \beta'_j \mathbf{X}_i}}{\sum_{k \in J} e^{\alpha_k + \beta'_k \mathbf{X}_i}} \forall j, k \in J$$

where $d_{ij} = 1$ if land use *j* is selected by agent *i*, and $d_{ij} = 0$ otherwise.

Defining $\tau \equiv \alpha_j \bigcup \beta_j \forall j \in J$, the log-likelihood function under the assumption that all land use decisions are accurately classified can be represented as:

$$\mathrm{LogL}(\tau) = \sum_{i=1}^{n} \sum_{j \in J} d_{ij} \ln \Pr_{ij}$$

Since in our analysis $J = \{AG, PC, GC, FA\}$

$$\mathrm{LogL}(\tau) = \sum_{i=1}^{N} \left(d_{i,AG} \ln \Pr_{i,AG} + d_{i,PC} \ln \Pr_{i,PC} + d_{i,GC} \ln \Pr_{i,GC} + d_{i,FA} \ln \Pr_{i,FA} \right)$$

As previously stated, agroforestry plantations normally have high biomass density which makes it practically impossible to identify fallowed parcels using only remotely sensed data. In the study area I consider that after years of low coffee prices farmers may not have had enough economic incentives to cultivate their plantations. Therefore, it is possible that some parcels classified in the Agroforestry category are in fact abandoned lands that should be categorized in the Fallow group. Additionally, some parcels that are classified as grass or corn may be incorrectly included in the Fallow category if as part of a rotational system those parcels were left uncultivated during some years to recover soil nutrients.

Considering that the proportion of parcels classified as AG may include a percentage of misclassified FA parcels, and that the latter class can have observations that should be in the GC category; we can represent the log likelihood function using missing information indicators to represent the hidden proportions. Let $d^{*}_{i,AG,FA}$ indicate the proportion of land use decisions in the AG category (branch) that are misclassified FA observations, and $d^{*}_{i,AG,AG}$ represent the proportion of accurately classified observations in the AG category satisfying the constraint $d^*_{i,AG,AG} + d^*_{i,AG,FA} = 1$; and similarly for $d^*_{i,FA,FA}$ and $d^*_{i,FA,GC}$. We can represent the log likelihood function as,

$$\mathrm{LogL}(\tau) = \sum_{i=1}^{N} \begin{pmatrix} d^*_{i,AG,AG} \ln \Pr_{i,AG,AG} + d^*_{i,AG,FA} \ln \Pr_{i,AG,FA} \\ + d_{i,PC} \ln \Pr_{i,PC} \\ + d_{i,GC} \ln \Pr_{i,GC,GC} \\ + d^*_{i,FA,FA} \ln \Pr_{i,FA,FA} + d^*_{i,FA,GC} \ln \Pr_{i,FA,GC} \end{pmatrix}$$

Unfortunately, since the proportions of land use observations that are correctly and incorrectly classified, $d^*_{i,j,k}$, is unknown we cannot identify the parameter estimates that maximize the log-likelihood function following the standard procedure. Nevertheless, we can replace the unknown proportions by their conditional expectations (Caudill, 2006).

$$E\left(d^*_{i,AG,AG} \middle| d^*_{i,AG}\right) = \frac{\exp\left(\alpha_{AG,AG} + \beta'_{AG,AG}\mathbf{X}_i\right)}{\exp\left(\alpha_{AG,AG} + \beta'_{AG,AG}\mathbf{X}_i\right) + \exp\left(\alpha_{AG,FA} + \beta'_{AG,FA}\mathbf{X}_i\right)}$$

$$E\left(d^*_{i,AG,FA} \middle| d^*_{i,AG}\right) = \frac{\exp\left(\alpha_{AG,FA} + \beta'_{AG,FA}\mathbf{X}_i\right)}{\exp\left(\alpha_{AG,AG} + \beta'_{AG,AG}\mathbf{X}_i\right) + \exp\left(\alpha_{AG,FA} + \beta'_{AG,FA}\mathbf{X}_i\right)}$$

$$E\left(d^*_{i,FA,FA} \middle| d^*_{i,FA}\right) = \frac{\exp\left(\alpha_{FA,FA} + \beta'_{FA,FA}\mathbf{X}_i\right)}{\exp\left(\alpha_{FA,FA} + \beta'_{FA,FA}\mathbf{X}_i\right) + \exp\left(\alpha_{FA,GC} + \beta'_{FA,GC}\mathbf{X}_i\right)}$$

$$E\left(d^*_{i,FA,GC} \middle| d^*_{i,FA}\right) = \frac{\exp\left(\alpha_{FA,GC} + \beta'_{FA,GC}\mathbf{X}_i\right)}{\exp\left(\alpha_{FA,FA} + \beta'_{FA,FA}\mathbf{X}_i\right) + \exp\left(\alpha_{FA,GC} + \beta'_{FA,GC}\mathbf{X}_i\right)}$$

where $E\left(d^*_{i,AG,AG} \middle| d^*_{i,AG}\right)$ indicates the conditional expectation that parcel i classified as AG is actually an AG parcel, and $E\left(d^*_{i,AG,FA} \middle| d^*_{i,AG}\right)$ represents the conditional expectation that a parcel classified as AG is in fact FA (the remaining conditional expectations have similar interpretations).

Defining

$$\varphi \equiv \exp\left[\alpha_{AG,AG} + \beta^{'}_{AG,AG}\mathbf{X}_i\right] + \exp\left[\alpha_{AG,FA} + \beta^{'}_{AG,FA}\mathbf{X}_i\right] + \exp\left[\alpha_{PC} + \beta^{'}_{PC}\mathbf{X}_i\right] + \\ \exp\left[\alpha_{GC} + \beta^{'}_{GC}\mathbf{X}_i\right] + \exp\left[\alpha_{FA,FA} + \beta^{'}_{FA,FA}\mathbf{X}_i\right] + \exp\left[\alpha_{FA,GC} + \beta^{'}_{FA,GC}\mathbf{X}_i\right]$$

The probabilities of observing agent i selecting each of the land use categories can be computed as

$$\Pr_{i,AG,j}\left(d_{i,AG,j} = 1 \middle| \mathbf{X}_i, \beta_{AG,j},\right) = \frac{e^{\alpha_{AG,j} + \beta^{'}_{AG,j}\mathbf{X}_i}}{\varphi} \; for\, j \; = \; AG,\; FA$$

$$\Pr_{i,FA,k}\left(d_{i,FA,k} = 1 \middle| \mathbf{X}_i, \beta_{FA,k}\right) = \frac{e^{\alpha_{FA,k} + \beta^{'}_{FA,k}\mathbf{X}_i}}{\varphi} \; for\; k \; = \; FA,\; GC$$

$$\Pr_{i,l}\left(d_{i,l} = 1 \middle| \mathbf{X}_i, \beta_l\right) = \frac{e^{\alpha_l + \beta^{'}_l\mathbf{X}_i}}{\varphi} \; for\; l \; = \; PC,\; GC$$

Under this modeling assumptions the log likelihood can be re-stated as,

$$\mathrm{LogL}(\tau) = \sum_{i=1}^{N}\left(\begin{array}{l} \mathrm{E}\left(d^{*}_{i,AG,AG} \middle| d^{*}_{i,AG}\right)\ln\dfrac{\exp\left(\alpha_{AG,AG} + \beta^{'}_{AG,AG}\mathbf{X}_i\right)}{\varphi} + \mathrm{E}\left(d^{*}_{i,AG,FA} \middle| d^{*}_{i,AG}\right)\ln\dfrac{\exp\left(\alpha_{AG,FA} + \beta^{'}_{AG,FA}\mathbf{X}_i\right)}{\varphi} \\ + d_{i,PC}\ln\dfrac{\exp\left(\alpha_{PC} + \beta^{'}_{PC}\mathbf{X}_i\right)}{\varphi} \\ + d_{i,GC}\ln\dfrac{\exp\left(\alpha_{GC} + \beta^{'}_{GC}\mathbf{X}_i\right)}{\varphi} \\ + \mathrm{E}\left(d^{*}_{i,FA,FA} \middle| d^{*}_{i,FA}\right)\ln\dfrac{\exp\left(\alpha_{FA,FA} + \beta^{'}_{FA,FA}\mathbf{X}_i\right)}{\varphi} + \mathrm{E}\left(d^{*}_{i,FA,GC} \middle| d^{*}_{i,FA}\right)\ln\dfrac{\exp\left(\alpha_{FA,GC} + \beta^{'}_{FA,FA}\mathbf{X}_i\right)}{\varphi} \end{array}\right)$$

To avoid identification problems the parameters associated with the stems of the branches that contain misclassified information need to be equivalent to the parameter estimates of the branches in which the parcels should be accurately classified. Therefore I set

$$\beta_{AG,AG} = \beta_{AG};$$
$$\beta_{AG,FA} = \beta_{FA};$$
$$\beta_{FA,FA} = \beta_{FA};$$
$$\beta_{FA,GC} = \beta_{GC}$$

Caudill (2006) highlights the relevance of the intercepts in the model since as $\alpha_{AG,FA} \to -\infty$ the probability of identifying fallowed parcels that are misclassified as agroforestry goes to zero. Similar reasoning applies when $\alpha_{FA,GC} \to -\infty$. To test that the LMNL model can be used to detect misclassified observations, we estimate profile likelihood confidence intervals (defined below) on those intercepts to tests that they are statistically different from $-\infty$. This will also constitute evidence that the related branch has misclassified parcels.

The procedure to determine the proportions of misclassified data and to compute the parameter estimates that maximize the likelihood function follows these steps:

1. Control for local maxima.

   Set a global solver or grid search algorithm to define vectors of initial values for the alternative specific parameters that will be estimated. This step is necessary since this LMNL modeling approach is related to a finite mixture model (Caudill et al., 2011) and during the computation of the parameter estimates we need to control for multiple local maxima of the likelihood function.

2. Expectation step.

   Use the observed data $\mathbf{X}_i$ and one of the vectors estimated in step 1 as initial values of the parameter estimates $\hat{\tau}^{(0)}$ to compute the conditional expectations of the misclassified and accurately classified land use proportions, $d_{i,jk}{}^{*}$.

3. Maximization step.

   Estimate the vector of parameters that maximize the likelihood function, $\hat{\tau}^{*}$, and the corresponding value of the likelihood function at that point $\mathrm{LogL}\left(\hat{\tau}^{*}\right)$.

4. Iterate between the expectation and maximization steps using $\hat{\tau}^{*}$ to update the conditional expectation of $d_{i,jk}{}^{*}$ and utilizing those values to re-compute $\hat{\tau}^{*}$ until the log-likelihood function convergences to a maximum value within a certain tolerance level (i.e., until $\mathrm{abs}\left(\mathrm{LogL}\left(\hat{\tau}^{*}\right)^{k} - \mathrm{LogL}\left(\hat{\tau}^{*}\right)^{k+1}\right) \leq \text{tolerance}$ value. Store $\mathrm{LogL}\left(\hat{\tau}^{*}\right)^{k+1}$ and $\hat{\tau}^{*}$.

5. Return to step 1 and repeat the process for a different vector of initial values $\hat{\tau}^{(0)}$ until exhausting the set of defined vectors in step 1.

6. Identify the $\hat{\tau}^{*}$ that produces the global maximum from the set of evaluated starting values.

4.4 Empirical model of land use decisions with misclassified data.

The dataset described in Chapter 2 is used to test for the presence of misclassified land use decisions and to evaluate the potential impact of such misclassification on the estimation results from a multinomial logit model. The modeling assumptions are similar to those described in Chapter 3, in the sense that I consider that agents' decisions are driven by changes in the relative expected net revenues across parcels. One difference with the mixed conditional logit model used in chapter 3 is that in this chapter for comparison purposes with commonly used models in the land use literature, I consider that revenue is a parcel specific characteristic. This can be justified by appealing to variations in unobserved components that impact yield quality/quantity or affect markets prices received by farmers in the study area. Additionally, instead of using the FA category for normalization purposes I normalize the coefficients of the PC since I am trying to assess whether or not there are misclassified observations within the FA group.

The model was coded within the Matlab environment using the global optimization toolbox to identify the parameter estimates that maximize the log-likelihood function. Overall the results indicate that around 11% of the observations contained in the sample are misclassified. Table 4.1 shows the parameter estimates ordered by branches and stems as well as the sum of the probabilities in each stem that indicates the number of observations accurately and inaccurately classified within each branch. At the stem level the first column shows the estimated coefficients that correspond to observations accurately classified as agroforestry as well as the number of observations that are contained in that subcategory. The second column contains the parameter estimates that correspond to misclassified FA observations within the AG branch.

Table 4.1. Latent multinomial logit model parameter estimates.

	BRANCHES									
	Agroforestry				Grass and Corn		Fallow			
	STEMS									
	Agroforestry		Fallow		Grass and Corn		Fallow		Grass and Corn	
Revenue	**0.1184**	***	**0.1549**	***	**0.1002**	***	**0.1549**	***	**0.1002**	**0.1184**
	9.0117		5.0535		5.2245		5.0535		5.2245	9.0117
Slope	**0.3549**	***	**9.2943**	***	**-0.2077**		**9.2943**	***	**-0.2077**	**0.3549**
	4.1403		5.3017		-2.1372		5.3017		-2.1372	4.1403
Distance to market	**0.3760**	*	**66.6943**	***	**0.9913**	***	**66.6943**	***	**0.9913**	**0.3760**
	1.9271		5.5119		5.7421		5.5119		5.7421	1.9271
Distance to road	**1.3163**	***	**46.7071**	***	**1.2370**	***	**46.7071**	***	**1.2370**	**1.3163**
	4.0694		5.4901		3.9381		5.4901		3.9381	4.0694
Poverty	**0.0835**		**-156.9003**		**-0.1447**		**-156.9003**		**-0.1447**	**0.0835**
	0.4365		-4.8884		-0.6993		-4.8884		-0.6993	0.4365
Soil texture	**0.0875**		**-180.2229**		**-0.5086**		**-180.2229**		**-0.5086**	**0.0875**
	0.4844		-6.1153		-3.1839		-6.1153		-3.1839	0.4844
Elevation	**5.1065**	***	**-221.7704**		**-2.4227**		**-221.7704**		**-2.4227**	**5.1065**
	7.1205		-4.8692		-3.5045		-4.8692		-3.5045	7.1205
Population	**-0.2368**		**25.3460**		**-0.4954**		**25.3460**		**-0.4954**	**-0.2368**
	-3.3417		4.7849		-6.8403		4.7849		-6.8403	-3.3417
Constant	**-3.3241**		**-9.0654**		**0.9822**	*	**-35.4194**		**0.9822**	**-3.3241**
	-5.9694		-0.1129		1.8411		0.0000		1.8411	-5.9694
$\sum_{i=1}^{n} \Pr_{ij}$	**547**		**52**		**536**		**0**		**108**	

Notes: The parcel specific coefficients of the Perennial Crops category were normalized to zero for model identification. The parameter estimates are shown in bold numbers; the numbers below each parameter estimate indicate the correspondent t-ratios. Significance codes: '***' significant at the 0.1% level; '**'significant at the 1% level; '*' significant at the 5% level; '.' Significant at the 10%.

The results indicate that 52 observations that are categorized as AG in the sample are in fact FA parcels. Those observations represent 8.7% of the parcels originally classified as AG. Recall that for identification purposes those slope coefficients are constrained to equal the corresponding estimates for the accurately classified observations within the FA branch. Similarly, the columns associated to each stem of the FA branch show the corresponding estimated coefficients and number of observations that should be classified in each stem. The results indicate that the procedure used to construct the FA category is suspect because all the observations in the FA branch-FA stem are considered misclassified by the LMNL procedure. In other words, the analysis provides evidence to argue that parcels that appear to be FA lands are in fact part of a rotational production system or parcels that continue under cultivation but that have not received maintenance activities during the time of data collection of the remotely sensed data utilized to produce the land use maps.

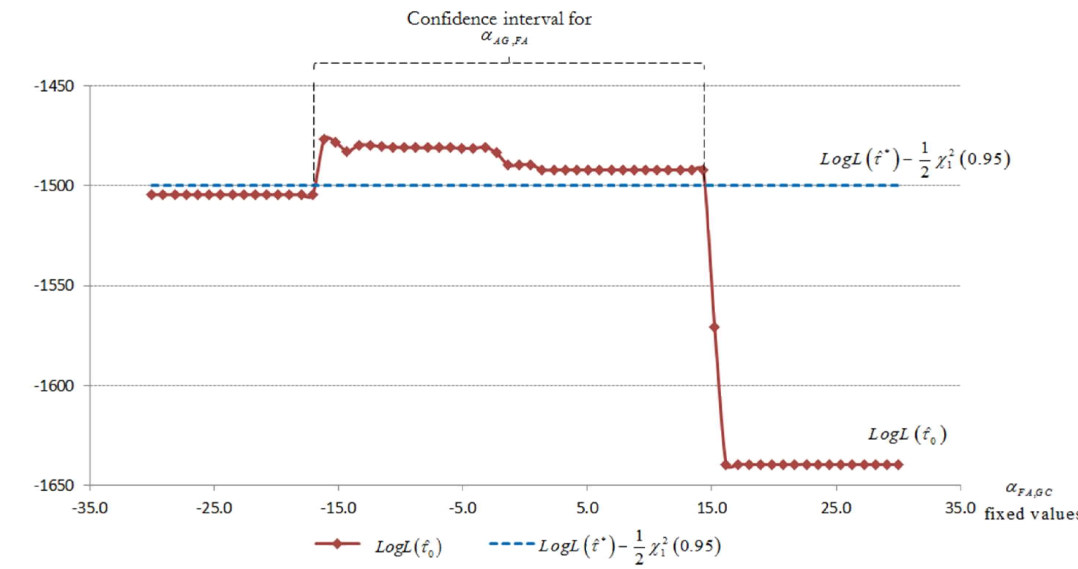
To test whether these results are statistically significant I compute profile likelihood confidence intervals for the intercepts $\alpha_{AG,FA}$ and $\alpha_{FA,GC}$ using a grid search procedure described by Stryhn and Christensen (2003). The lower and upper bounds of a profile likelihood confidence interval for a parameter $\alpha_{j,k}$ are the ones that satisfy the equation $LogL(\hat{\tau}^*) - \frac{1}{2}\chi_1^2(0.95) \leq LogL(\hat{\tau}_0)$, where $\hat{\tau}^*$ are the maximum likelihood parameter estimates (MLE) of τ, $\chi_1^2(0.95)$ indicates the 95% quantile of a chi-squared distribution with one degree of freedom, and $\hat{\tau}_0$ is a vector that contains the MLE of τ obtained after setting the parameter of interest to a fixed value x (i.e., $\alpha_{j,k} = x$), and treating the remaining parameters in the model as nuisance parameters. The aforementioned grid search procedure identifies the values of x for which the inequality

$LogL\left(\hat{\tau}^{*}\right)-\frac{1}{2}\chi_{1}^{2}\left(0.95\right)\leq LogL\left(\hat{\tau}_{0}\right)$ holds. The profile likelihood confidence interval for $\alpha_{AG,FA}$ is [-17.1, 15.2] and for $\alpha_{FA,GC}$ is [-1.29, 0.77]. Clearly the corresponding lower bounds are bounded away from $-\infty$, which provides evidence that the number of misclassified observations is statistically significant greater than zero. Figure 4.1 shows the profile likelihood confidence intervals for both parameters of interest.

A depiction of the differences between the land use proportions in the sample data and the percentages estimated with the LMNL model is presented in Figure 4.2. The results indicate that the AG category is overrepresented in the sample during all the study period due to the presence of misclassified observations. On the other hand, the GC category is underrepresented in the sample since it should contain all the observations categorized as FA in the sample dataset. For the same reason, the FA category appears to be overrepresented during all the period of analysis. A potential explanation for this finding is that small-landowners that rely primarily on household labor are less likely to abandon their plantations (Albers et al., 2006) specially if the current land use provides means to satisfy household subsistence constraints.

To analyze the impacts of misclassified observations on the magnitudes and directions of the parameter estimates I use the original sample dataset and the reconstructed sample based on the LMNL analysis to estimate a multinomial logit model of land use decisions. Table 4.2 shows the estimated coefficients, significance levels and standard errors.

Figure 4.1. Profile likelihood confidence intervals for $\alpha_{AG,FA}$ and $\alpha_{FA,GC}$.



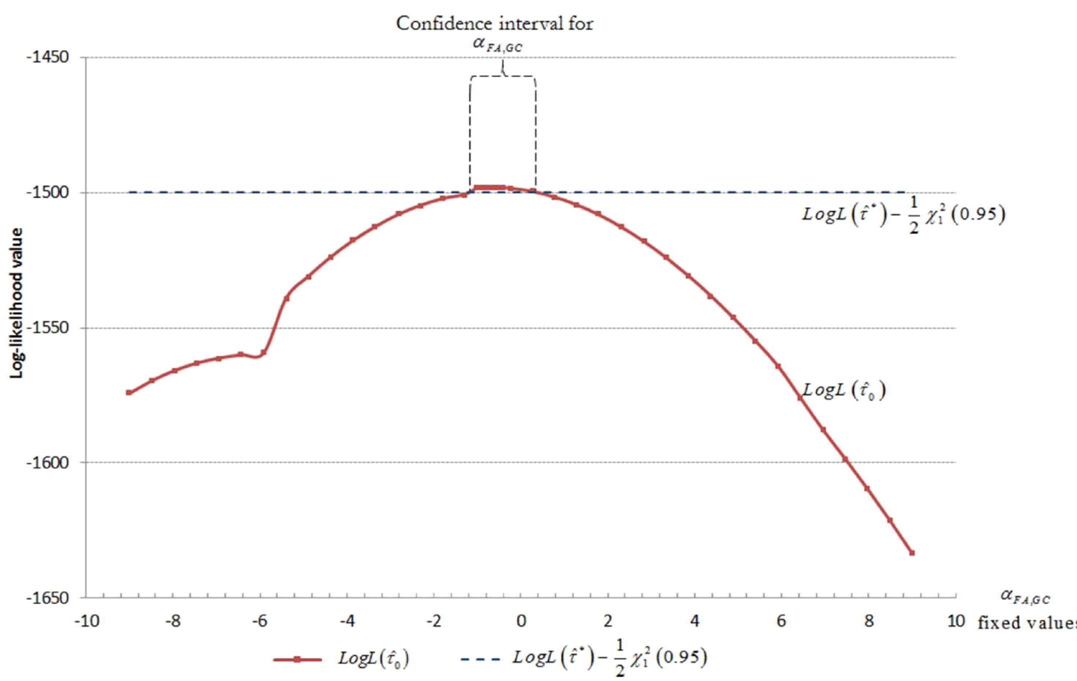
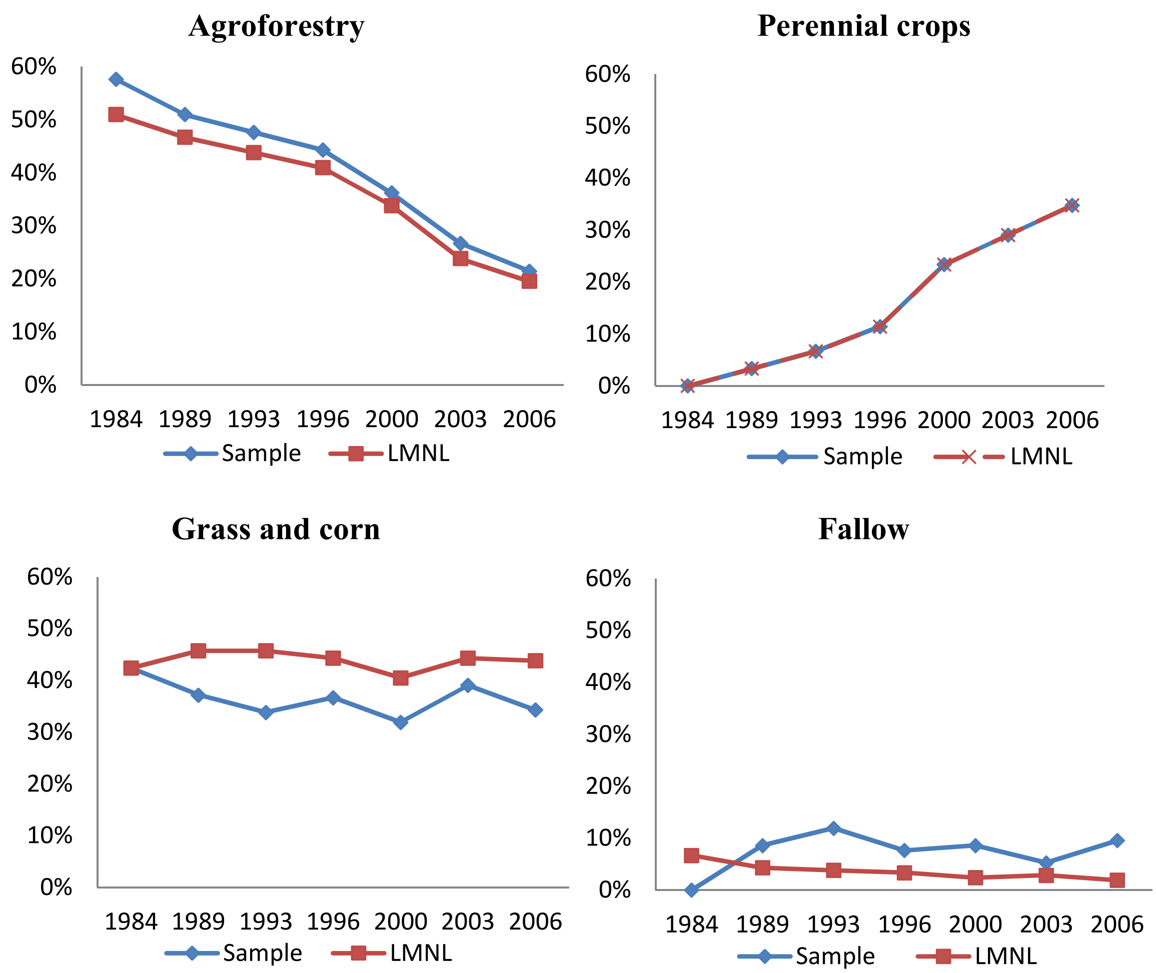


Figure 4.2. Land use proportions in the sample data and estimated proportions using the LMNL model.



Overall the significance levels and values of the AG and GC parameter estimates are similar in the analysis of the two sample datasets. The values of the coefficients associated to the FA category appear to be significantly different in magnitude and in some cases the signs change using the LMNL-based sample. Given the significant reconfiguration of the observations contained in the FA category the differences in the corresponding parameter estimates was expected. Since the coefficients of parcel specific variables do not directly allow us to infer how changes in the independent variables affect the probability of observing any of the land uses considered in the analysis, I computed the change in the probability of observing land use *j* in parcel *i* resulting from a marginal change in the magnitude of each of the independent *k* variables included in the multinomial land use model. The individual calculations were averaged across parcels and land uses and the results are shown in Table 4.3. In general, most of the marginal effects estimated with the two datasets have the expected directions (See chapter 2 for a detailed description of the expected marginal effects). According to the analysis there is statistical evidence to argue that parcels with higher degrees of slope will be more likely to be used for agroforestry production, and areas with low slope are preferred for cornfields or grasslands. The average marginal effects of the distance from a parcel to the nearest markets are statistically significant and have the expected signs. The probability of observing cash crops (AG or PC) reduces as the distance to a market increases. On the other hand, the likelihood of an agent selecting the GC or FA category increases as the distance to the nearest market increases, which is consistent with the intuition that if a parcel is located far away from a market transportation costs may reduce the profitability of some of the land uses reducing the choice set to subsistence crops (such as corn), land uses that require a large area (such a cattle ranching activities), or land abandonment. A similar explanation applies to the average marginal effects of the variable measuring the distance from a parcel to the nearest road. Although, the marginal effect for this variable has the expected direction only in the estimation output corresponding to the reconfigured dataset.

Table 4.2. Multinomial logit parameter estimates using the original sample data and the reconfigured sample data generated with the LMNL model.

		Original sample data			Redistributed sample data			
		Estimate (A)	Std. Error		Estimate (B)	Std. Error		Difference A - B
Constant	AG	-3.3421	0.5539	***	-3.3805	0.5630	***	0.0384
	GC	1.2527	0.5424	*	1.1672	0.5365	*	0.0855
	FA	0.5276	0.8060		2.1882	2366.36		-1.6605
Slope	AG	0.3789	0.0841	***	0.3505	0.0848	***	0.0284
	GC	-0.2557	0.0946	**	-0.1956	0.0937	*	-0.0601
	FA	-0.1833	0.1389		3.1100	0.6253	***	-3.2934
Distance to market	AG	0.5782	0.1685	***	0.3214	0.1784	.	0.2567
	GC	0.8627	0.1713	***	1.0540	0.1805	***	-0.1913
	FA	0.9732	0.2086	***	21.1698	4.2179	***	-20.1966
Distance to road	AG	1.4016	0.3185	***	1.3174	0.3267	***	0.0841
	GC	1.2489	0.3119	***	1.2305	0.3118	***	0.0184
	FA	1.0312	0.3852	**	16.1232	3.1221	***	-15.0920
Poverty index	AG	-0.0154	0.2328		0.1164	0.2320		-0.1318
	GC	-0.2802	0.2344		-0.1733	0.2267		-0.1069
	FA	0.7179	0.3157	*	-45.4773	9.7809	***	46.1952
Soil texture	AG	0.0896	0.1642		0.0728	0.1656		0.0168
	GC	-0.5437	0.1655	**	-0.5168	0.1610	**	-0.0268
	FA	-0.3428	0.2533		-62.3399	2366.38		61.9971
Elevation	AG	4.6620	0.7143	***	5.2670	0.7357	***	-0.6051
	GC	-1.9084	0.7206	**	-2.6215	0.7295	***	0.7130
	FA	-2.7255	1.1215	*	-64.3404	14.3039	***	61.6149
Population	AG	-0.2124	0.0664	**	-0.2149	0.0674	**	0.0025
	GC	-0.4901	0.0739	***	-0.5067	0.0734	***	0.0165
	FA	-0.5253	0.1291	***	6.7800	1.5575	***	-7.3054
Revenue	AG	0.1090	0.0127	***	0.1196	0.0132	***	-0.0105
	GC	0.0852	0.0188	***	0.1013	0.0192	***	-0.0161
	FA	-0.0123	0.0114		0.1846	0.0384	***	-0.1969
Log-Likelihood:		-1492.4			-1187.4			
McFadden R^2:		0.16417			0.2939			

Notes: The parcel specific coefficients of the Perennial Crops category were normalized to zero for model identification. Significance codes: '***' significant at the 0.1% level; '**'significant at the 1% level; '*' significant at the 5% level; '.' Significant at the 10%.

Table 4.3. Average marginal effects.

			Original sample data (A)		Redistributed sample (B)		Difference (B - A)
		Expected sign	Estimate	Standard Deviation	Estimate	Standard Deviation	
Slope	AG	+	0.103	0.032	0.077	0.056	-0.026
	GC	-	-0.085	0.028	-0.089	0.067	-0.003
	FA	+ -	-0.011	0.009	0.020	0.099	0.031
	PC	-	-0.007	0.021	-0.009	0.021	-0.002
Distance to market	AG	-	-0.007	0.037	-0.116	0.275	-0.109
	GC	+	0.070	0.033	0.071	0.442	0.000
	FA	+	0.021	0.015	0.132	0.647	0.111
	PC	-	-0.084	0.051	-0.087	0.067	-0.003
Distance to road	AG	-	0.103	0.063	0.049	0.216	-0.054
	GC	+	0.052	0.055	0.004	0.325	-0.048
	FA	+	-0.005	0.012	0.097	0.473	0.102
	PC	-	-0.150	0.089	-0.150	0.093	0.000
Poverty index	AG	-	0.013	0.014	0.144	0.601	0.131
	GC	+	-0.076	0.038	0.130	0.946	0.207
	FA	+	0.056	0.044	-0.294	1.440	-0.350
	PC	-	0.008	0.008	0.020	0.098	0.012
Soil texture	AG	+	0.084	0.032	0.215	0.821	0.131
	GC	-	-0.104	0.031	0.139	1.296	0.242
	FA	+-	-0.006	0.009	-0.402	1.967	-0.396
	PC	+	0.025	0.025	0.048	0.135	0.023
Elevation	AG	+	1.140	0.354	1.378	0.936	0.238
	GC	+-	-0.793	0.303	-0.846	1.490	-0.053
	FA	+-	-0.208	0.137	-0.419	2.048	-0.211
	PC	-	-0.140	0.239	-0.113	0.325	0.027
Population	AG	+	0.025	0.021	0.007	0.097	-0.019
	GC	-	-0.054	0.020	-0.092	0.145	-0.039
	FA	-	-0.013	0.009	0.046	0.227	0.059
	PC	+	0.041	0.026	0.039	0.033	-0.002
Revenue	AG	+	0.011	0.005	0.009	0.006	-0.003
	GC	+	0.005	0.006	0.003	0.006	-0.001
	FA	+	-0.006	0.005	0.001	0.003	0.006
	PC	+	-0.010	0.006	-0.013	0.008	-0.002

Expected direction codes: ‘ + ‘ indicates that a positive marginal effect is expected, ‘ – ‘ indicates that a negative marginal effect is expected, ‘ + - ’ indicates that the marginal effects can go in either direction. See chapter 2 for a detailed description of the expected direction of the marginal effects.

The results corresponding to the poverty index are statistically significant only for the FA category. I would expect that richer areas have higher probability of selecting cash crops although this does not correspond to the results from the LMNL dataset. None of the parameter estimates for the variable that controls for soil texture are statistically significant, and besides considering that the GC category should be more likely to be selected in areas with finer soil texture, it is difficult to argue about the expected direction of the marginal effects on the remaining categories. All the parameter estimates of the elevation variable are statistically significant and the directions of the marginal effects of the AG and PC categories are consistent with the agroecological requirements of the crops in those land use classes. Since corn and grass can be produced in parcels located at different elevation gradients, the direction of the marginal effects could go in either direction depending on the location of the parcels in the dataset. The results for the original and modified sample data indicate an inverse relationship between elevation and the probability of observing GC and FA. The parameter estimates corresponding to the population variable are statistically significant and the marginal effects have the expected signs indicating that higher population density may increase the probability of observing labor intensive land uses and vice versa.

Perhaps the most relevant results after the sample data reconfiguration are related to the statistical significance of the estimated coefficients of the revenue variable and the signs of the corresponding marginal effects. The parameter estimates computed with the original dataset, that contains misclassified data, are statistically significant at the 0.1% level for the AG and GC category and the marginal effects have theoretical consistent signs. Nevertheless, the marginal effects of changes in revenue on the probability of an agent selecting the FA or PC categories indicate a counterintuitive direction. Those inconsistencies appear to be partially corrected with the reconfiguration of the dataset based on the results of the LMNL model. Specifically, the sign of the revenue related marginal effects for the FA category have the expected sign although the multinomial logit model still cannot provide parameter estimates that explain how despite a continuous decrease in the revenue of the PC category, some of the farmers decided to transition to that land use category.

4.5 Conclusion.

Given the limited availability of historical high resolution remotely sensed data, land use change analyses are often restricted to the study of transitions between a reduced set of choices. In some cases coarse datasets are enough to accomplish relevant research objectives, for instance in the study of deforestation processes. Nevertheless, in most of the spatially explicit land use analysis coarse land use classifications are implemented as a mechanism to reduce classification errors. Unfortunately, even in land use datasets composed by a reduced number of categories, classification errors are still a potential modeling issue. Consider for example an analysis that uses only two categories, forested and agricultural lands, to study deforestation drivers in a particular region. In this case it is possible that some of the observations classified as forested areas are in fact fallowed parcels devoted to agricultural production, or even grasslands that have not received weed control activities during the time of data collection of the remotely sensed data. Unfortunately, those types of classification errors are difficult to reduce using only pixel-based algorithms, particularly if the available land use information is part of a cross-sectional dataset, or time series data with many years of separation between the observed periods.

As an alternative to reduce classification errors, in this chapter I implement a post-classification procedure to identify misclassified land use data that cannot be detected using pixel-based classification algorithms. The Latent Multinomial Logit methodology has been implemented in several contexts to detect misclassified categorical data (Caudill et al., 2005, 2011; Caudill & Mixon Jr., 2005; Caudill, 2006) but to our knowledge it has not been applied in the land use change literature. The analysis implemented in this chapter is based on land use information generated with remotely sensed data collected during seven points in time throughout the period 1984 -2006, with a maximum separation of five years between observations. The data correspond to land use transitions observed in a Mexican coffee growing region in which relatively high rates of tree canopy removal were observed as a result of the clearing of shade-grown coffee plantations. I analyze land use dynamics between agroforestry parcels, perennial crops, grass and corn, and land abandonment. The category corresponding to fallow lands was constructed analyzing the sequence of land use decisions observed in each parcel and setting a parcel equal to fallow when the land use oscillated between GC or PC and AG within a period of at most six years.

The implementation of the LMNL model provides statistical evidence to argue that the procedure used to construct the FA category, while reasonable and objectively defensible, fails to recognize that temporary increases in biomass that appear to indicate a change in the corresponding land use classification to AG, may be a result of a production system that requires land abandonment as a mechanism to recover soil productivity; or simply an indication that the parcel has not been maintained during the time in which the remotely sensed data in that region were collected. On the other hand, the results indicate that the LMNL procedure can be used to identify parcels within the AG category that have a high likelihood of being fallowed only using the information produced with pixel-based classification algorithms without making any assumptions about the land use sequence followed by each landowner. With regard to the impact on the values and magnitudes of the parameter estimates and marginal effects of the data statistically identified as misclassified, we can observe that in general the reclassification of the parcels based on the LMNL model increases the magnitudes of the marginal effects in the theoretically expected direction. Particularly, the marginal effect of changes in revenue associated with the FA category becomes statistically significant with the theoretically expected direction.

Unfortunately, the available land use information does not allow us to assess the accuracy of the LMNL model. I do not have data to confirm that the parcels identified as misclassified land use observations are in fact devoted to a land use that is different to the one detected by the classification algorithms. Simulated data or a sample of land use decisions validated with interviews of landowners could help to test the performance of the approach used here. Additionally, we can test the argument that fallowed parcels in the PC category are not a common situation by expanding the LMNL model to let it estimate whether or not there are some parcels in that category that statistically appear to be misclassified lands.

Chapter 5

Structural estimation of stochastic dynamic agent-based land use decisions.

5.1. Introduction

The main objective of this paper is to investigate the bias associated with the use of a reduced form static (myopic) model in the analysis of a dynamic land use decision process. Similar investigations have been conducted previously in the broader resource economics literature but apparently have gone largely unnoticed by land use change modelers. Baerenklau and Provencher (2005) analyze the implications of using a static model to study recreation site choices that presumably are the outcome of a dynamic process. They conclude that even if a reduced form static model provides a good fit to the observed data it fails to produce unbiased and theoretically consistent estimates. Hicks and Schnier (2006) reach a similar conclusion after analyzing site choices for commercial fishing. Despite these related efforts, the vast majority of the agent-based land use change literature still overlooks the fact that landowners tend to be dynamic forward-looking agents who consider the future consequences of their current decisions and who maximize uncertain payoffs over multiple time periods.

The number of studies that use dynamic frameworks to analyze land use decisions appears to be very limited, and usually not focused on the estimation of the structural components of the decision making process. The Forest and Agriculture Sector Optimization Model (Adams et al., 1996) uses a deterministic dynamic programming approach to identify optimal land use allocations among competing activities but does not undertake parameter estimation. Schatzki (2003) observes that land use models that fail to incorporate uncertain returns and conversion costs may overestimate the price responsiveness of decision makers and produce inaccurate policy assessments. The author shows that option value models can provide a better analysis of conversion decisions, but his approach assumes land use choices are irreversible which can introduce significant bias into the model predictions if decisions are, in fact, reversible (Baerenklau and Knapp, 2007). Song, Zhao and Swinton (2011) expand Schatzki's (2003) model to allow for the possibility of reversible decisions but focus on the impact of alternative stochastic return processes on land use choices, rather than estimation of the underlying decision problem.

In perhaps the most comprehensive study to date, De Pinto and Nelson (2008) use a dynamic discrete choice model that incorporates agent expectations about future prices, short term irreversibility of decisions, sunk costs and parcel specific characteristics (e.g., slope, elevation, soil quality) to analyze landscape transformation in a Central American country. That paper constitutes the first attempt to estimate the structural parameters that describe the land use decision processes in a particular geographical area, and to contrast those parameters with those generated by static models. The authors find that the dynamic model produces more accurate predictions than the static models, and the parameter estimates appear to be statistically different, although no formal analysis is presented to validate that observation. The De Pinto and Nelson (2008) analysis constitutes a relevant contribution to the land use literature, but nevertheless exhibits some shortcomings. The discretization of the state space is very coarse, and instead of constructing a discrete state space with equidistant points distributed within a range with lower and upper bounds defined around the observed minimum and maximum values of the stochastic variables, the authors use the observed values as the reference discrete states, i.e., assume that future realizations of those variables are centered at one of the historically observed values. The data is similarly coarse in the temporal dimension, and makes the common assumption that land use changes occur in the observation year. This assumption fails to consider that the observed land use at time t may have been in fact selected at time $t - n$, as a result of the state of the world and subjective expectations at that time. Last, the estimation method sacrifices efficiency for computational speed. The present work aims to address these issues and provide clearer insight into the implications of static modeling of dynamic land use change processes.

The empirical application focuses on explaining landscape dynamics in a Mexican coffee growing region during the period 1984-2006. In that time a substantial amount of tree canopy was lost as farmers moved out of shade coffee and into other crops like corn, citrus and banana. One of the motivations of this chapter is to model how commodity price fluctuations impact canopy loss and how price policies might be used to mitigate such losses. Supervised classification methods are implemented on seven satellite images that cover the study period with an average separation between observations of 3.7 years. To develop a panel dataset, the resulting classification is coupled with topographic and socioeconomic data that affect land use productivity/returns (e.g., slope, distance to markets, soil type, population, distance to nearest communities, output prices, conversion costs). To reduce the effects of spatial autocorrelation in the analysis systematic random sampling is used to identify a subset of parcels representative of the landscape configuration and land use changes observed in the study region. Additionally, the available remotely sensed data is used along with the physiological constraints of the crops to estimate land uses for the periods in which land use observations could not be obtained.

Agent-based land use decisions are modeled as dependent on future returns expectations, conversion costs, and parcel characteristics within a framework that allows for the possibility that land use decisions can be reversed. I analyze output prices received by farmers in the study area and approximate the subjective price expectation process followed by those agents. A multinomial logit structure is imposed on each choice occasion and Rust's (1987) efficient nested fixed-point algorithm is implemented to solve the associated Bellman equation. To accelerate the numerical estimation, spline interpolation is used in the analysis. Estimating the vector of structural parameters that best describes the observed land use choices requires solving the stochastic dynamic optimization problem many times "inside" a maximum likelihood parameter estimation routine.

5.2 Stochastic dynamic discrete choice land use model

As stated in previous chapters, land use analysis based on discrete choice random utility models typically assume that the state of different socioeconomic and environmental factors at a particular time $t = 1, ..., T$, for $T \leq \infty$, determines the choice set of land uses available to decision maker i at different t periods, C_{it}, with $i \in \eta = \{1, 2, ...n\}$. For a specific geographic region an analyst can identify a finite set of feasible land uses $J = \bigcup_{i \in \eta}^{t \in \tau} C_{it}$. I assume that landowners are forward-looking agents who observe, at the beginning of each period, the realizations of the variables that determine current payoffs, and use that information to update their estimates of future values for the stochastic variables present in their decision making process. I assume that each agent selects the land use $j \in J$ that maximizes the present value of his current and expected future flow of payoffs. According to Rust (1994) for Markovian discrete decision processes Blackwell's theorem implies that with perfect information on all the state variables that determine agent's choices, an adequate modeling approach would be able to accurately predict the observed behavioral process. Nevertheless, since only a subset $\mathbf{x}$ of the state variables is observed by the analyst, I assume that deviations from the model predictions are explained by state variables, ε_j, that determine the payoffs of each land use $j \in J$ and that are only observed by each landowner.

Recognizing that the land use decisions considered in this study are reversible, I model land use choices as part of an infinite horizon optimization problem followed by agents that try to maximize a time-homogeneous utility function in which the unobservable components enter the utility function in an additively separable (AS) way that can be represented as

$$u_j(\mathbf{x}) + \varepsilon_j \ \forall j \in J \tag{1}$$

with an additional assumption that agents use a time invariant transition probability function to estimate future states of the stochastic land use drivers, and a constant $\beta \in [0,1)$ to discount future payoffs. These assumptions correspond to a stationary Markovian structure of land use decisions and imply that future expected payoffs depend exclusively on the current state of the land use drivers and the current land use decision. In other words, if an agent at time *t* and at time *t* + *k* reaches a situation in which the land use drivers are in state *s,* the optimal decision rule and value function should be equivalent (Rust, 1994). Therefore, we can represent the agent's value function associated to the Bellman equation as

$$V(\mathbf{x},\ \varepsilon) = \max_{j \in J}\left[u_j(\mathbf{x}) + \varepsilon_j + \beta\left\{V(\mathbf{x}',\ \varepsilon')\,p\left(\mathbf{x}',\ \varepsilon'\middle|\mathbf{x},\ \varepsilon, j\right)\right\}\right] \tag{2}$$

where $p\left(\mathbf{x}',\ \varepsilon'\middle|\mathbf{x},\ \varepsilon, j\right)$ represents a transition probability equation used by landowners to estimate the likelihood of observing the realization of the stochastic components of the observable and unobservable land use drivers $(\mathbf{x}',\ \varepsilon')$ conditional on their current realizations $(\mathbf{x},\ \varepsilon)$, and land use during the current period. Under the assumption of conditional independence (CI) the transition equation can be represented as

$$p\left(\mathbf{x}',\varepsilon'\middle|\mathbf{x},\varepsilon\right) = q\left(\varepsilon'\middle|\mathbf{x}'\right)\pi\left(\mathbf{x}'\middle|\mathbf{x},j\right)$$

This decomposition implies that the probability of a land use driver assuming a particular realization during the next period only depends on the current period realization of that state variable, and on the agent's land use choice. This also implies that any serial dependence between ε' and ε is contained in the next period realization of the observable state variables $\mathbf{x}'$ (Rust, 1987). More importantly, the AS and CI assumptions allow us to estimate a system of conditional choice probabilities, $\Pr_j\left(d_j = 1 \middle| \mathbf{x}\right)$ where $d_j = 1$ if land use j is selected and $d_j = 0$ otherwise, that can be used to estimate the structural parameters of the discrete choice dynamic problem(Rust 1994).

Therefore under the AS and CI assumptions the value function (2) can be rewritten as,

$$V\left(\mathbf{x},\ \varepsilon\right) = \max_{j \in J}\left[u_j\left(\mathbf{x}\right) + \varepsilon_j + \beta E_{\varepsilon'} V\left(\mathbf{x}',\ \varepsilon'\right) \pi\left(\mathbf{x}' \middle| \mathbf{x}, j\right)\right] \tag{3}$$

Following Card (2008) we can define

$$\begin{aligned} \bar{V}\left(\mathbf{x}\right) &\equiv E_\varepsilon V\left(\mathbf{x}, \varepsilon\right) \\ &\equiv E_\varepsilon\left[\max_{j \in J}\left[u_j\left(\mathbf{x}\right) + \varepsilon_j + \beta \bar{V}\left(\mathbf{x}'\right)\ \pi\left(\mathbf{x}' \middle| \mathbf{x}, j\right)\right]\right] \end{aligned} \tag{4}$$

Representing the expected payoff to land use j as,

$$v_j\left(\mathbf{x}\right) \equiv u_j\left(\mathbf{x}\right) + \beta \bar{V}\left(\mathbf{x}'\right)\ \pi\left(\mathbf{x}' \middle| \mathbf{x}, j\right) \tag{5}$$

we can rewrite equation 4 in a way that resembles the model specification of a static discrete choice problem

$$\bar{V}\left(\mathbf{x}\right) = E_\varepsilon\left[\max_{j \in J}\left[v_j\left(\mathbf{x}\right) + \varepsilon_j\right]\right]$$

Assuming that the unobservable state variables are i.i.d. standard extreme value type I deviates

$$\bar{V}(\mathbf{x}) = E_{\varepsilon}\left[\max_{j\in J}\left[v_j(\mathbf{x})+\varepsilon_j\right]\right] = \log\left[\sum_j \exp\left(v_j(\mathbf{x})\right)\right]+\gamma \tag{6}$$

where $\gamma = 0.577215$ is the Euler-Mascheroni constant.

The probability of an agent selecting land use j can be computed using the commonly known expression

$$\begin{aligned}\Pr_j\left(d_j = 1|\mathbf{x}\right) &= \frac{\exp\left(v_j(\mathbf{x})\right)}{\sum_{l\in J}\exp\left(v_l(\mathbf{x})\right)} \\ &= \frac{\exp\left[u_j(\mathbf{x})+\beta\bar{V}(\mathbf{x}')\ \pi\left(\mathbf{x}'|\mathbf{x},j\right)\right]}{\sum_{l\in J}\exp\left[u_l(\mathbf{x})+\beta\bar{V}(\mathbf{x}')\ \pi\left(\mathbf{x}'|\mathbf{x},l\right)\right]}\end{aligned} \tag{7}$$

As mentioned in previous chapters, the payoffs of the different land uses available to the decision maker are determined by several factors that include a parcel's characteristics, an agent's characteristics, the regional climate, governmental policies, etc. Following the land use literature we can model the systematic component of the utility function, $u_j(\mathbf{x})$, as a linear-in-parameter specification of the observable state variables. Furthermore, we can decompose the vector $\mathbf{x}$ into variables that represent parcel specific characteristics that do not change through time and parcel specific stochastic variables. Therefore, we can define the current period payoff to land use j from the observable components in equation 7 as,

$$u_j(\mathbf{x}) \equiv \omega_j{}'\mathbf{s}_i + \alpha_j{}'\mathbf{k}_i$$

where $\mathbf{k}_i$ is a vector of parcel i specific variables, $\mathbf{s}_i$ is a vector of stochastic variables, and ω_j and α_j are structural parameters of the decision making process associated with each land use in the choice set.

With a set of observed land use decisions made by $i = 1, ..., N$ agents, at different time periods $t = 0, ..., T$, we can compute the parameters, ω_j and α_j that maximize the likelihood function and characterize more accurately the decision making process followed by landowners in the study region:

$$L\left(\omega, \alpha \middle| \mathbf{s}, \mathbf{k}, d_{ij}\right) = \prod_{i=1}^{N} \prod_{j} \prod_{t=0}^{T} \mathrm{Pr}_{ijt}^{d_{ijt}}$$

or in log-likelihood functional form,

$$\begin{aligned} \log L &= \sum_{i=1}^{N} \sum_{j} \sum_{t=0}^{T} d_{ijt} \log \mathrm{Pr}_{ijt} \\ &= \sum_{i=1}^{N} \sum_{j} \sum_{t=0}^{T} d_{ijt} \log \frac{\exp\left[v_j(\mathbf{x})\right]}{\sum_{l} \exp\left[v_l(\mathbf{x})\right]} \quad \forall \;\; l, j \in J \end{aligned}$$

5.3 Empirical application

As stated in the description of the study region, from an economic, ecological and social perspective, agroforestry production is a significant activity in Mexico. Small-scale farmers across the country depend upon shade grown crops, with coffee being the leader both in terms of cultivated land area and value of production. It has been estimated that around 3 million Mexican people depend on coffee-related activities and that approximately 90% of the coffee-cultivated area lays under diversified shade providing forest-like environmental services (Escamilla Prado, 2007). The modeling efforts implemented with the dataset described in chapter 2, so far consists of a multinomial logit model implemented by Ellis et al. (2010), and Baerenklau et al. (2012), the mixed multinomial – conditional logit described in chapter 3 and the Latent multinomial logit model presented in chapter 4. None of those models have the methodological structure to appropriately model inter-temporal aspects associated to forward-looking behavior. Recognizing the limitations of those analyses this paper focuses on implementing an analysis of the structural components of the dynamic decision-making process followed by landowners in the low lands of Atzalan, Veracruz, Mexico. The following subsections present a description of specific modifications made to the dataset described in chapter 2, and additional state variables generated to implement Rust's (1987) nested-fixed point algorithm.

5.3.1 Land use choice set.

The choice set of land uses available to landowners in the study region is defined by the categories described in chapter 2, Agroforestry (AG), Perennial crops (PC), Grass and Corn (GC), and Fallow lands (FA). In this chapter, I use information from the sample of 210 parcels of land use decisions obtained with remotely sensed data collected during 1984, 1989, 1993, 1996, 2000, and 2006. The observations in the FA category correspond to the parcels that were identified as abandoned lands after analyzing the sequence of land use transitions derived from the classification based on remotely sensed data[3].

5.3.2 Transition probability matrices for the price indices

Recall that the AG price index is defined as the average rural price received by coffee growers at the State level. The PC price index is a weighted average price per-ton according to the area harvested for each product in this category. The GC price index is a per-hectare weighted price according to the average productivity of grasslands and cornfields. The FA price index is based on the yearly minimum wage for construction workers, which is a proxy of the alternative income that farmers can obtain if they decide to let their land fallow and work off farm. To reduce the dimensionality of the state space we normalize the price indices using as a reference the FA category. This will allow the estimation of the dynamic model with only three stochastic variables although the normalization will limit our analysis to a multinomial (rather than conditional) logit specification.

In this chapter I assume that landowners at the beginning of the period observe the realizations of each of the price indices for the land uses under analysis, and update their subjective price expectations to estimate the profitability of changing to an alternative land use or continuing the current land use for at least one additional period. Since one of the assumptions required to implement the nested fixed-point algorithm is that the transition probability function is stationary, I needed to ensure that the time series of the

[3] The reclassification generated with the latent multinomial logit model described in chapter 4 is not used in the dynamic analysis described in this chapter. I decided not to use that classification since it has not been validated, but mainly because a more robust implementation of the latent multinomial logit model would include the dynamic components of the land allocation process.

observed prices do not exhibit structural breaks that induce changes in the subjective price expectation process used by decision makers during the period of analysis. Particular attention was paid to the analysis of the AG price index, since the international coffee market suffered a reconfiguration in 1989 at the end of the International Coffee Agreement, moving from an export quota system to an unregulated export market. To test for structural breaks in the time series of the normalized price indices I use the *strucchange* R package (Zeileis, Leisch, Hornik, & Kleiber, 2002) to implement generalized fluctuation tests and F-tests (see Kuan, Leisch, and Hornik 2000). In general, these analyses do not indicate statistical evidence to reject the null hypothesis of structural stability for the time series considered in the model.

Given that price indices are considered stochastic components of the land use model under analysis, it is necessary to estimate the associated transition probability functions that are used by agents to elicit future price expectations. Arguably the most reliable method to do this involves collection of experimental survey data from decision-makers. Because such surveys are costly and time consuming and thus beyond the scope of this thesis, I instead base my approach on findings from a recent study by Hill (2010) that does this for Ugandan coffee growers. Using survey data, Hill (2010) demonstrates that recent coffee prices are one of the main determinants of farmers' future expectations. Therefore, consistent with these findings as well as the requirements of my estimation framework, I assume a first-order auto-regressive process for the normalized coffee price. Furthermore, because coffee production in the study region was one of the most important sources of income during the 1950s – 1990s, I assume that farmers use a similar auto-regressive structure to estimate prices for other crops as well. Although it is possible that different processes could be driving price expectations for each farmer in the study region, to make the empirical model tractable, and for comparison purposes with the work by De Pinto and Nelson (2008), I apply a common set of parameter estimates to all farmers. Estimation results are shown in Table 5.1.

Table 5.1. First order autoregressive parameters of the price indices used in the estimation of the transition probability functions.

Parameter	AG	PC	GC
Unconditional mean	0.2306	0.0814	0.3968
Autocorrelation coefficient	0.6742	0.5050	0.8213
Standard deviation of the shocks	0.0758	0.0215	0.0425

To estimate the associated transition probability matrices, I follow a two-step process. First, I discretize each normalized price index into a finite set of states $\mathbf{Price}_s = \left\{ \text{Price}_{s,1}, \text{Price}_{s,2}, \ldots, \text{Price}_{s,n} \right\}$ for $s = \text{AG, PC, GC}$ and compute the n x n marginal probability transition matrix $\prod_s = \left[\pi_{ij} \right]_{i,j\,=\,1,\ldots,n}$, where π_{ij} represents the probability of $\text{Price}_{s,\,i}$ moving next period to $\text{Price}_{s,\,j}$ $\forall$ $i, j = 1,\ldots,n$; $\pi_{ij} \geq 0$ and $\sum_j \pi_{ij} = 1$ $\forall$ i. To populate the $\prod_s$ matrixes I implement Tauchen and Hussey's (1991) approximation algorithm of AR(1) processes with Markov chains. In the second step, under the assumption of price independence we compute the n^3 x n^3 joint transition probability matrix as the Kronecker product of the marginal probability transition matrices $\prod_{join} = \Pi_{AG} \otimes \Pi_{PC} \otimes \Pi_{GC}$.

5.3.3 Age of the plantations

In a dynamic forward-looking framework, the age of the land uses is a key variable since it determines the expected yields for the AG and PC categories, and consequently the expected revenues for those land uses. Recall that in the original dataset the largest separation between land use observations occurs between 1973 and 1984, and that to define a starting value for the age variable, I identified 210 parcels that did not change land use during those periods. In that sample it is assumed that the minimum age of each land use *i* at the beginning of 1984 is 11 years, $age_{i,\ 1984} = 11$. It is likely that most of the land uses in that sample in 1984 are actually older than 11 years, but since at that age it is assumed that all the land use categories have reached maturity, the assumption that $age_{i,\ 1984} = 11$ seems an adequate starting point to construct a yearly dataset of land use ages during all the study period (1984 – 2006).

To estimate ages after 1984, recall that the dataset contains land use observations for the periods at which the satellite images were collected (1984, 1989, 1993, 1996, 2000, 2006). Unfortunately, after land use is changed in a particular parcel the new land use observed during the next observation period does not necessarily correspond to a one-year old plantation. For instance, if the land use of parcel *i* in 1984 is AG and in 1989 is PC, the land use change may have happened at any date after the data collection time of 1984, and the age of the PC plantation in that parcel at the beginning of 1989 may, for modeling purposes, range from one to five years. Let (*t, t'*) = {(1984,1989); (1989,1993); (1993,1996); (1996,2000); (2000,2003); (2003,2006)}, indicate the observed land use at the beginning and end of the intervals for which we have land use information. To construct a yearly dataset of age values I assume that if the land use at *t'* is the same as the one at *t*, then it is not economically rational and physiologically possible to change to a different land use during the intervening period and then back to the observed land use at *t*, especially in the case of perennial crops. On the other hand, if $\mathrm{d}_{ijt} = 0$ and $\mathrm{d}_{ijt'} = 1$ the land use transition to land use *j* might have happen any time after *t* during the years for which we do not have land use observations or at time *t'*. Letting $age_{ij,\,t}$ represent the age of land use *j* in parcel *i* at the beginning of year *t*, and $age_{ij,\,t'}$ indicate the age at the beginning of next year *t'* for which there is land use information, we can consider that the age associated to the observed land uses in all the parcels can be represented as,

$$age_{ij,\,t'} = \begin{cases} 0 & \text{if } \mathrm{d}_{ijt} = 0 \\ z \in \{1,2,\ldots,t'-t\} & \text{if } \mathrm{d}_{ijt} = 0 \text{ and } \mathrm{d}_{ijt'} = 1 \\ age_{ij,\,t} + t' - t & \text{if } \mathrm{d}_{ijt} = 1 \text{ and } \mathrm{d}_{ijt'} = 1 \end{cases} \qquad \forall\, j \in J$$

With perfect information we will be able to identify the land use transition time for each parcel with land use change, determine the value of z, and incorporate that information in the computation of the likelihood function. Given data limitations I investigate the sensitivity of three procedures to assign a transition time to all the observed land use changes. The first procedure, which is generally assumed in most of the spatially explicit land use models, considers that the observed land use change was made in the observation period, which implies that $age_{ij,t'} = 1$ if $d_{ijt} = 0$, and $d_{ijt'} = 1$. The second procedure reconstructs a yearly dataset of land use age values assumes that the observed land use change happened during the year following period *t*, (i.e., $age_{ij,t'} = t' - t$ if $d_{ijt} = 0$, and $d_{ijt'} = 1$). The third procedure assumes that the land use transition happened in the middle of the transition period *t* – *t'* using the average $age_{ij,t'} = {t'-t}/{2}$ if $d_{ijt} = 0$, and $d_{ijt'} = 1$). I tested these procedures and find that the results are not significantly different, arguably because the time separation between the observed land use decisions is at most 5 years, which is not common in the land use literature. I decided to follow the assumption that the observed land use change occurred in the middle of the transition periods to construct a yearly dataset of land use age information. In addition, I impose an upper bound of 25 years in order to reduce the size of the state space under the assumption that the returns from older plantations become indistinguishable at this age[4].

5.3.4 Parcel specific characteristics

Given the computational requirements to estimate the maximum likelihood parameter estimates associated with the land use problem under analysis, I decided to use a reduced set of parcel specific variables. During the model calibration process, the variables that account for elevation, distance to nearest markets and poverty level were observed to be statistically relevant to the dynamic analysis, whereas slope, population,

[4] An alternative to estimating the age variable is to consider it as a "missing completely at random" (MCAR) data problem generated by the lack of information about the specific time in which land use change occurred. Under this assumption, the EM algorithm and multiple imputation, then deletion method (MID) (Enders, 2010; Young & Johnson, 2010) can be applied, albeit at a significantly additional computational cost.

distance from each parcel to the nearest road and soil texture were not and were thus omitted from the dynamic analysis. Table 5.2 shows a summary of the parcel-specific variables that are used to estimate the dynamic model.

Table 5.2. Summary statistics parcel specific variables used in the dynamic model

Variable	Description	Mean	Min	Max
Elevation	Meters above sea level	353.95	85	726.40
Poverty	Poverty index	0.316	-0.798	2.109
Distance to nearest market	Distance from each parcel to nearest market (m)	14.36	2.93	35.52

Summary statistics are computed for all parcels during the period 1984-2006.

5.4 Results and discussion

To compute the maximum likelihood parameter estimates (MLE) of the structural land use model, the nested fixed point algorithm was implemented using the student version of Matlab R2012a. For a discretized set of values of the discount factor $\beta \in [0,1)$, the Global Optimization toolbox was used to implement first a scatter-search algorithm to produce vectors of starting coefficient values; and then a gradient-based constrained nonlinear optimization procedure to estimate the local optimum within the neighborhood of each vector of starting values. The local solutions were stored and the values of the log-likelihood function were compared to identify the global optimum in the set of tested starting coefficient values. Once a MLE was associated with each of the initial β values, a derivative-free unconstrained procedure was implemented to refine the parameter estimates. This last step was used because during the model calibration process it was detected that, for some coefficient values, the gradient-based procedures apparently reached relatively flat regions of the log-likelihood function, thus making it difficult for the solver to identify a direction to move to increase the log-likelihood value. The derivative-free procedures helped to ensure that no further improvements, within the neighborhood of a tolerance value, could be made to the values of the parameter estimates to increase the value of the objective function.

To expedite the computational procedure, as is common in dynamic models with continuous stochastic variables, the state space corresponding to each of the normalized price indexes of the AG, PC, and GC categories was discretized into 12 equidistant points distributed through a range with lower and upper bounds defined by the minimum and maximum values observed during the observation period. The corresponding transition probability matrixes generated with those sets of discretized values is shown in Tables 5.3 – 5.5. Additionally, to estimate value functions associated with observed price realizations that are different from the discrete points used in the analysis, I use cubic spline interpolation. Furthermore, the main computational scripts and functions used in the analysis were configured to run in parallel to use more efficiently the available computing resources.

5.4.1 Sensitivity of the model to different discount factors

To initiate the computational estimation the discount factor, β, was discretized into the set of points $\beta = \{0.05, 0.25, 0.5, 0.75, 0.85, 0.95\}$. The grid search procedure used to estimate the discount factor indicates that a value around 0.5 provides relatively good estimates of the observed land use proportions and of the inter-temporal land use trajectories. Therefore, the discrete state space for the variable discount factor was further

Table 5.3. Transition probability matrix
of the discretized state space corresponding to the AG normalized price index

Discrete points	***0.09323***	***0.12826***	***0.16328***	***0.19831***	***0.23333***	***0.26836***	***0.30338***	***0.33841***	***0.37343***	***0.40846***	***0.44348***	***0.47851***	**Sum**
0.09323	0.11218	0.37913	0.35409	0.13094	0.02187	0.00173	0.00006	0.00000	0.00000	0.00000	0.00000	0.00000	1
0.12826	0.01801	0.16850	0.37507	0.30838	0.11030	0.01828	0.00141	0.00005	0.00000	0.00000	0.00000	0.00000	1
0.16328	0.00220	0.04910	0.22920	0.37249	0.25507	0.07964	0.01154	0.00075	0.00002	0.00000	0.00000	0.00000	1
0.19831	0.00020	0.01016	0.09374	0.28513	0.35488	0.19839	0.05123	0.00598	0.00029	0.00001	0.00000	0.00000	1
0.23333	0.00001	0.00153	0.02697	0.14913	0.32805	0.31951	0.14311	0.02910	0.00251	0.00008	0.00000	0.00000	1
0.26836	0.00000	0.00017	0.00553	0.05478	0.20996	0.35135	0.26920	0.09404	0.01415	0.00080	0.00001	0.00000	1
0.30338	0.00000	0.00001	0.00080	0.01415	0.09404	0.26920	0.35135	0.20996	0.05478	0.00553	0.00017	0.00000	1
0.33841	0.00000	0.00000	0.00008	0.00251	0.02910	0.14311	0.31951	0.32805	0.14913	0.02697	0.00153	0.00001	1
0.37343	0.00000	0.00000	0.00001	0.00029	0.00598	0.05123	0.19839	0.35488	0.28513	0.09374	0.01016	0.00020	1
0.40846	0.00000	0.00000	0.00000	0.00002	0.00075	0.01154	0.07964	0.25507	0.37249	0.22920	0.04910	0.00220	1
0.44348	0.00000	0.00000	0.00000	0.00000	0.00005	0.00141	0.01828	0.11030	0.30838	0.37507	0.16850	0.01801	1
0.47851	0.00000	0.00000	0.00000	0.00000	0.00000	0.00006	0.00173	0.02187	0.13094	0.35409	0.37913	0.11218	1

Table 5.4. Transition probability matrix
of the discretized state space corresponding to the PC normalized price index

Discrete points	***0.04854***	***0.05614***	***0.06375***	***0.07135***	***0.07895***	***0.08655***	***0.09416***	***0.10176***	***0.10936***	***0.11696***	***0.12457***	***0.13217***	**Sum**
0.04854	0.01371	0.14541	0.36025	0.32685	0.12841	0.02333	0.00197	0.00007	0.00000	0.00000	0.00000	0.00000	1
0.05614	0.00209	0.04745	0.22532	0.37192	0.25849	0.08188	0.01204	0.00080	0.00002	0.00000	0.00000	0.00000	1
0.06375	0.00031	0.01347	0.11138	0.30628	0.34625	0.17624	0.04147	0.00441	0.00020	0.00000	0.00000	0.00000	1
0.07135	0.00004	0.00330	0.04548	0.20003	0.35377	0.27855	0.10103	0.01661	0.00115	0.00003	0.00000	0.00000	1
0.07895	0.00000	0.00069	0.01553	0.10687	0.28957	0.34558	0.18935	0.04717	0.00502	0.00020	0.00000	0.00000	1
0.08655	0.00000	0.00012	0.00443	0.04714	0.19361	0.34656	0.28389	0.10608	0.01710	0.00104	0.00002	0.00000	1
0.09416	0.00000	0.00002	0.00104	0.01710	0.10608	0.28389	0.34656	0.19361	0.04714	0.00443	0.00012	0.00000	1
0.10176	0.00000	0.00000	0.00020	0.00502	0.04717	0.18935	0.34558	0.28957	0.10687	0.01553	0.00069	0.00000	1
0.10936	0.00000	0.00000	0.00003	0.00115	0.01661	0.10103	0.27855	0.35377	0.20003	0.04548	0.00330	0.00004	1
0.11696	0.00000	0.00000	0.00000	0.00020	0.00441	0.04147	0.17624	0.34625	0.30628	0.11138	0.01347	0.00031	1
0.12457	0.00000	0.00000	0.00000	0.00002	0.00080	0.01204	0.08188	0.25849	0.37192	0.22532	0.04745	0.00209	1
0.13217	0.00000	0.00000	0.00000	0.00000	0.00007	0.00197	0.02333	0.12841	0.32685	0.36025	0.14541	0.01371	1

Table 5.5. Transition probability matrix
of the discretized state space corresponding to the GC normalized price index

Discrete points	*0.24610*	*0.27442*	*0.30274*	*0.33105*	*0.35937*	*0.38769*	*0.41601*	*0.44433*	*0.47265*	*0.50096*	*0.52928*	*0.55760*	Sum
0.24610	0.35188	0.43990	0.17594	0.02982	0.00237	0.00009	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	1
0.27442	0.07675	0.33177	0.38222	0.17145	0.03442	0.00325	0.00014	0.00000	0.00000	0.00000	0.00000	0.00000	1
0.30274	0.00953	0.11868	0.33704	0.34674	0.15355	0.03135	0.00298	0.00013	0.00000	0.00000	0.00000	0.00000	1
0.33105	0.00073	0.02396	0.15609	0.34460	0.31600	0.13115	0.02520	0.00218	0.00008	0.00000	0.00000	0.00000	1
0.35937	0.00003	0.00291	0.04180	0.19108	0.35066	0.28623	0.10759	0.01833	0.00132	0.00003	0.00000	0.00000	1
0.38769	0.00000	0.00022	0.00669	0.06214	0.22426	0.35393	0.25588	0.08430	0.01194	0.00064	0.00001	0.00000	1
0.41601	0.00000	0.00001	0.00064	0.01194	0.08430	0.25588	0.35393	0.22426	0.06214	0.00669	0.00022	0.00000	1
0.44433	0.00000	0.00000	0.00003	0.00132	0.01833	0.10759	0.28623	0.35066	0.19108	0.04180	0.00291	0.00003	1
0.47265	0.00000	0.00000	0.00000	0.00008	0.00218	0.02520	0.13115	0.31600	0.34460	0.15609	0.02396	0.00073	1
0.50096	0.00000	0.00000	0.00000	0.00000	0.00013	0.00298	0.03135	0.15355	0.34674	0.33704	0.11868	0.00953	1
0.52928	0.00000	0.00000	0.00000	0.00000	0.00000	0.00014	0.00325	0.03442	0.17145	0.38222	0.33177	0.07675	1
0.55760	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00009	0.00237	0.02982	0.17594	0.43990	0.35188	1

refined around 0.5. Table 5.6. shows the maximum likelihood parameter estimates corresponding to some values of the expanded discretized state space for the discount factor. All but one of the signs of the parameter estimates remain unchanged across all the tested discount factor levels, the only changes are in the magnitudes of the coefficient estimates. Nevertheless, since the model is based on a multinomial logit procedure, the signs of the coefficients do not necessarily indicate the direction of the marginal effects.

Before proceeding with the estimation of the marginal effects it was necessary to identify the discount factor that provides the best approximation to the decision making process in the study region. Typically, the discount factor that produces the highest log-likelihood value is the one selected as the "true" value used by agents to discount future payoffs. Consistent with observations on market interest rates, previous studies generally have estimated relatively high values of the discount factor, for instance De Pinto and Nelson (De Pinto & Nelson, 2008) find that a β value of 88.5% maximizes the likelihood function associated to their dynamic model of land use change. However, despite testing alternative model specifications, I find that lower discount factors —as low as 5%— consistently produce higher likelihood values for the present analysis (figure 5.1). Because a 5% discount factor seems unrealistically low, I applied three alternative selection criteria for the "true" discount factor. First, I analyzed how accurately the model predicts land uses at the parcel level within the sample. The results also indicate a negative relationship between the per-parcel prediction accuracy, and the discount factor (figure 5.2). Second, following Keane and Wolpin (2009), I analyzed how accurately the model predicts land uses at the parcel level out of sample. To do this, I use observed land uses in 2011 derived from Google Earth imagery for the sample parcels. Figure 5.3 shows that discount factors ranging from 5% to 75% can predict the observed land uses in 2011 with an accuracy of around 82% without significant variation. The third, method focuses

Table 5.6. Maximum likelihood parameter estimates

Discount factor		5%	25%	45%	50%	57%	75%	85%	95%
Revenue	AGF	6.8409	6.1905	5.2345	5.3644	5.2383	4.6461	4.9333	3.9098
	PC	17.5313	13.7274	9.6098	9.7398	8.0886	6.1316	5.3308	2.4381
	GC	-0.7963	-1.8549	-2.1703	-2.3847	-2.4524	-8.6514	-9.0646	-10.4484
Intercept	AGF	-1.4843	-1.5304	-1.8057	-2.0392	-2.2148	-4.6034	-4.9291	-5.6746
	PC	-1.1761	-1.2194	-1.8500	-2.0835	-2.1615	-6.3249	-6.6509	-6.8180
	GC	2.5626	3.6150	3.8647	3.9945	3.8612	8.5354	8.2832	7.4046
Elevation	AGF	3.9908	3.4868	5.0454	5.1753	4.7987	14.4728	11.6728	11.2371
	PC	5.1013	3.9738	6.4255	6.5554	6.2144	18.4690	15.5218	16.2920
	GC	0.4177	-0.6338	-0.9282	-1.1617	-1.2286	-0.1473	-0.1536	-0.1651
Distance to market	AGF	-0.3181	-0.3705	-0.4737	-0.6034	-0.6240	-1.4868	-1.5901	-1.8421
	PC	-1.1639	-1.0512	-1.2713	-1.5048	-1.5465	-2.9611	-3.3425	-3.4118
	GC	-0.0796	-0.0706	0.0067	-0.0660	-0.0681	-0.2754	-0.2927	-0.2872
Poverty index	AGF	-0.8622	-0.1057	-0.1863	-0.3160	-0.3139	-2.7023	-2.7513	-2.7264
	PC	-1.2323	-0.7091	-0.9519	-1.1854	-1.1989	-2.0393	-2.0740	-2.0558
	GC	-0.8003	-0.4151	-0.4174	-0.5471	-0.5778	-1.3513	-1.3651	-1.7872
Log likelihood		-2631	-2807	-3202	-3368	-3680	-7089	-10041	-19882

Figure 5.1. Discount factor and Log Likelihood values

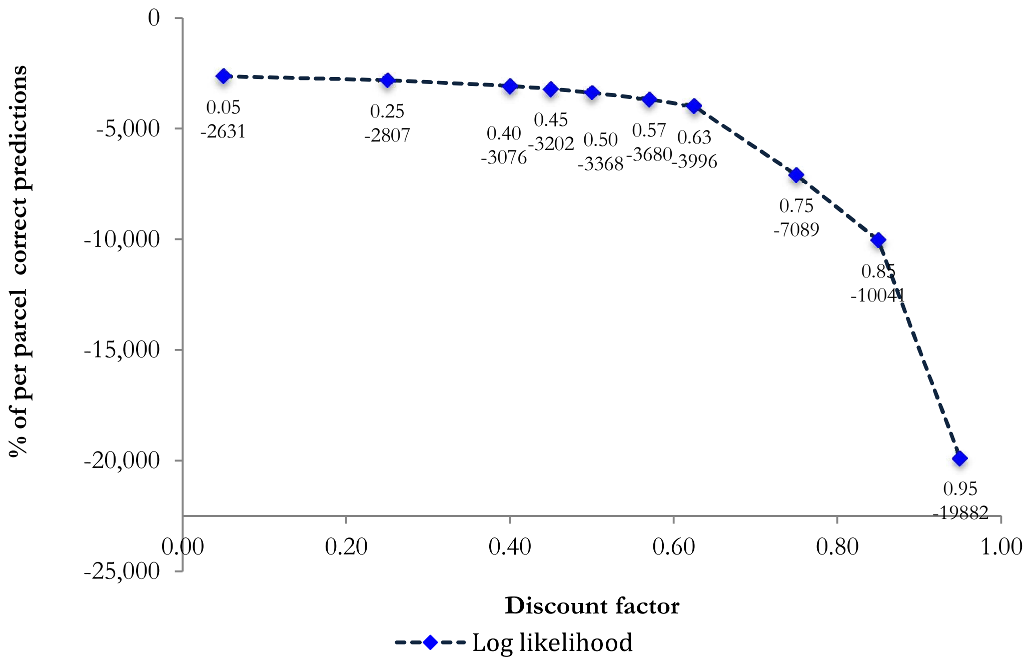


Figure 5.2. Percentage of per-parcel accurate predictions.

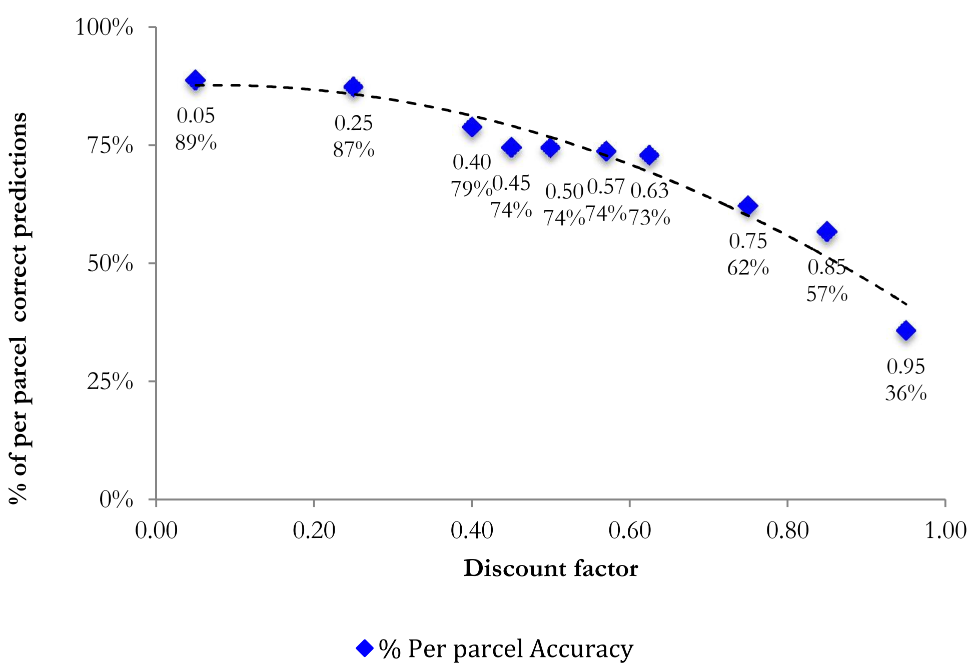
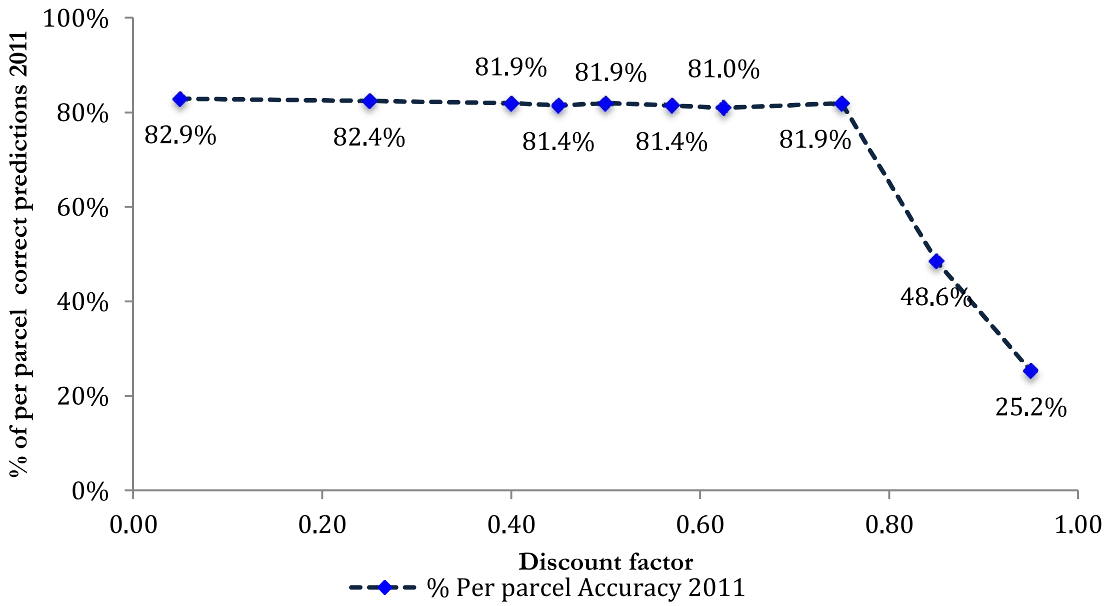


Figure 5.3. Out-of sample percentage of per-parcel accurate predictions.



on analyzing the estimated land use proportion during all the study period and at each year for which there available land use information. The results indicate that a dynamic model with a 57% performs well in the estimation of the aggregate land use proportions during the study period (see figure 5.4), as well as during the years for which there is land use data (see figure 5.5). Additionally, in table 5.6 we observe that despite the relatively good predictive performance of the model at a discount factor of 5%, it overestimates by 10.8% the percentage of parcels devoted to the GC category, which represents an overestimation of 29.7% of the parcels allocated to this land use category. Additionally, at that discount factor the model does not predict any of the parcels identified as fallowed lands. On the other hand, we can also observe in figure 5.4 and table 5.6 that as the discount factor approaches 100% the model tends to under predict the proportion of land in the AG, PC, and GC categories, and over predicts the percentage of land that is fallowed.

Overall, the results from the first two applied alternative procedures to detect the "true" discount factor appear to indicate that a discrete choice model with a discount factor equal to or close to zero is a good instrument to describe the behavioral process followed by agents in the study region. Furthermore, such a discount rate would imply that agents do not take into account future payoffs when making land use decisions, and would call into question the dynamic modeling approach. However, it is necessary to proceed with caution before drawing such a conclusion. As emphasized by Provencher and Baerenklau (2005), measures of fit are problematic tools for assessing whether an underlying decision process is dynamic. For example, under some circumstances even a low order polynomial can be used to accurately predict decisions that are inherently dynamic, but such a "reduced form" model lacks any insight into the structure of the decision making process. Instead, Provencher and Baerenklau (2005) argue that other information about the decision making process should be brought to bear on this question rather than relying on the data to answer it. Therefore, although a near-zero discount factor could be justifiable for annual crop production, particularly in regions where agents face subsistence constraints, it would be difficult to defend that farmers with coffee, banana or citrus plantations do not care about future flows of net revenue during their decision making process. Following this reasoning, a discount factor of 57% is used in the remainder of this analysis since it not only performs relatively well under each of the three selection criteria but also is more consistent with forward-looking behavior.

Figure 5.4. Estimated and observed use proportions at different levels of the discount factor.

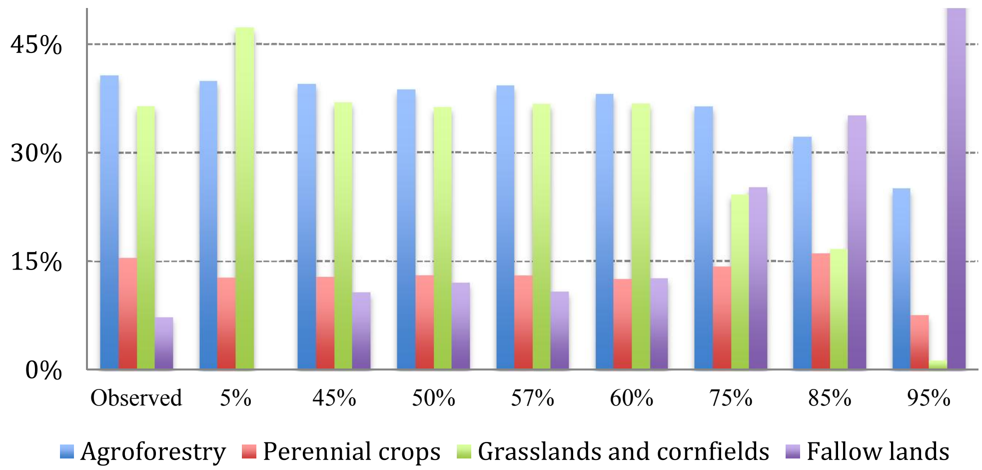


Figure 5.5 Observed versus estimated land use in the sample data at different discount factor values

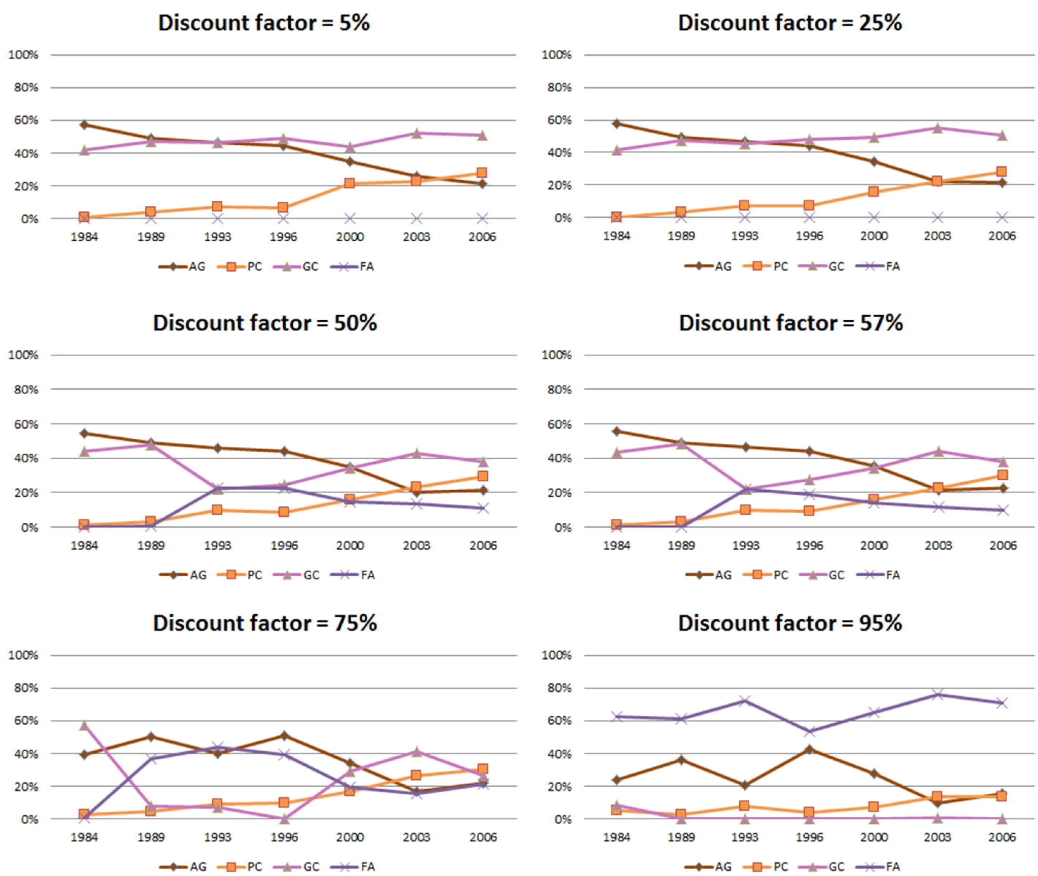


Table 5.7. Observed and predicted land use proportions at different levels of the discount factor.

	Observed land use proportions	Estimated land use proportions at different Discount factor values									
		5%	25%	45%	50%	57%	60%	63%	75%	85%	95%
Agroforestry	40.7%	39.9%	39.5%	39.5%	38.6%	39.3%	38.0%	38.1%	36.3%	32.2%	25.1%
Perennial crops	15.5%	12.8%	12.2%	12.9%	13.1%	13.1%	12.6%	12.4%	14.3%	16.1%	7.6%
Grasslands and cornfields	36.5%	47.3%	48.4%	36.9%	36.3%	36.7%	36.7%	36.8%	24.1%	16.7%	1.4%
Fallow lands	7.3%	0.0%	0.0%	10.7%	12.0%	10.9%	12.7%	12.7%	25.2%	35.1%	65.9%
Deviations from observed											
Agroforestry		-0.7%	-1.2%	-1.2%	-2.0%	-1.4%	-2.7%	-2.6%	-4.4%	-8.5%	-15.6%
Perennial crops		-2.7%	-3.3%	-2.7%	-2.4%	-2.4%	-2.9%	-3.1%	-1.2%	0.5%	-7.9%
Grasslands and cornfields		10.8%	11.9%	0.5%	-0.2%	0.3%	0.3%	0.3%	-12.3%	-19.8%	-35.1%
Fallow lands		-7.3%	-7.3%	3.4%	4.7%	3.5%	5.3%	5.3%	17.9%	27.8%	58.6%
Sum of absolute deviations											
		21.6%	23.8%	7.8%	9.4%	7.6%	11.2%	11.3%	35.8%	56.6%	117.1%

5.4.2 Comparison of the magnitudes and directions of marginal effects between static and dynamic models.

For comparison purposes, with the reduced set of explanatory variables considered in the dynamic model, I estimated a mixed conditional-multinomial model that uses non-normalized moving average prices to approximate agents' price expectations, and the average expected yield for a new plantation during the first 25 years of production to approximate farmers' expectations on plantation productivity to construct the revenue indexes (see chapter 3 for a description of this type of model). Additionally, similar modeling assumptions but using Fallow normalized prices were used to implement a standard multinomial logit model. Given that the parameter estimates produced by multinomial logit based models, both under static and dynamic frameworks, do not indicate the direction of the marginal effects, I computed the change in the probability of observing *AG, PC, GC* or *FA* in each sample parcel resulting from a marginal change in the magnitude of each of the independent variables included in the dynamic model. Table 5.8 shows the estimated marginal effects, statistical significance of the parameters estimates[5], and expected directions of the marginal effects produced with the aforementioned static models, and with the two versions of the dynamic model defined at discount factors of 5% and 57%.

[5] Since the values and signs of the parcel specific parameter estimates do not allow us to make direct comparisons or inferences, those results are not shown in the table

Table 5.8. Marginal effects estimated with static and dynamic models.

			Mixed Multinomial Conditional logit		Standard Multinomial logit		Dynamic 05%		Dynamic 57%	
		Expected sign	Estimate		Estimate		Estimate		Estimate	
Revenue	AG	+	**0.0047**	.	**0.2648**		**0.2593**	***	**0.7634**	
	PC	+	-0.0133	***	-0.0866	***	**0.8615**	***	**0.7745**	***
	GC	+	-0.0108	*	-0.2480	**	-0.9888		-1.5326	
	FA	+	-0.0059	***	**0.0698**		-0.1321		-0.0053	
Elevation	AG	+	**1.1352**	***	**1.1341**	***	**0.1579**	***	**0.6447**	*
	PC	-	0.0774	***	0.0793	***	0.2117	***	0.5143	**
	GC	+-	-0.9866		-0.9889		-0.2820		-1.1507	
	FA	+-	-0.2261		-0.2245		-0.0876		-0.0083	
Distance to market	AG	-	**-0.0141**	**	**-0.0117**	***	**-0.0049**		**-0.0164**	
	PC	-	**-0.1105**	***	**-0.1133**	***	**-0.0535**		**-0.1375**	***
	GC	+	**0.0995**		**0.1003**		**0.0438**		**0.1517**	
	FA	+	**0.0251**		**0.0247**		**0.0146**		**0.0022**	
Poverty index	AG	-	**-0.0190**	***	**-0.0181**	***	**-0.0061**		0.0790	
	PC	-	**-0.0645**	***	**-0.0658**	***	**-0.0269**		**-0.0852**	
	GC	+	**0.0167**	***	**0.0169**	***	-0.0239		**0.0029**	
	FA	+	**0.0669**		**0.0670**		**0.0569**		**0.0034**	

Notes: 1) Bold numbers indicate that the estimated marginal effects have the expected signs.

2) The parcel specific coefficients of the Perennial Crops category were normalized to zero for model identification. Significance codes: '***' significant at the 0.1% level; '**'significant at the 1% level; '*' significant at the 5% level; '.' Significant at the 10%.

3) Expected direction codes: ' + ' indicates that a positive marginal effect is expected, ' – ' indicates that a negative marginal effect is expected, ' + - ' indicates that the marginal effects can go in either direction. See chapter 2 for a detailed description of the expected direction of the marginal effects.

Since the static and dynamic models have different modeling assumptions about the process followed by decision makers to estimate the profitability of the available land uses, it is not possible to draw comparisons between the magnitudes of the marginal effects. Nevertheless, the directions of the marginal effects allow us to compare the performance of the implemented modeling approaches. Perhaps the most relevant finding is that the results from the dynamic model, at the tested discount factor levels, indicate theoretically consistent directions in the marginal effects for the two cash crops categories, AG and PC, by contrast the static models produce counterintuitive estimations for the PC category. The results for the GC and FA categories do not coincide with the expected direction, which may be due to the fact that the proportions of those land use categories remained relatively stable during the period of analysis. This, in turn, may be related to household consumption constraints, particularly for parcels devoted to corn production.

On the other hand we can observe that the estimated marginal effects for the parcel specific variables are very similar across the static and dynamic models. This is not entirely unexpected since the variables considered in the analysis are spatially heterogeneous but assumed to be time invariant at least during the period of analysis. All the implemented models indicate that higher altitudes increase the likelihood of observing coffee based agroforests, which as mentioned in previous chapters is consistent with the requirements to produce high quality coffee. The results also indicate that land use categories comprised of cash-crops are more likely to be in parcels located closer to markets, and that land uses that require large land areas, e.g., grasslands, or that may be devoted to producing goods for household consumption, e.g., cornfields, are more likely to be found relatively far from markets. With regard to the relationship between poverty levels and land use decisions, all the models indicate that citrus or banana plantations, the components of the PC category, are more likely to be found in areas with higher welfare status.

5.4.3 Welfare effects from a price floor policy simulation

The complications associated with the estimation of structural discrete choice dynamic programing models are usually justified in terms of their usefulness in the evaluation of counterfactual scenarios or policies (Keane & Wolpin, 2009). In the context of land use modeling, structural models can be used to analyze the impact of improving access to financial credit for small-holder farmers, price subsidies or price floors, technology assistance to improve parcel productivity, reduction in travel time to commercialize yields through road improvements, etc. Programs such as these were implemented by the Mexican government in response to the dramatic decline in coffee prices received by farmers during the 1990's and early 2000's that reached levels that were too low to cover production costs. There were two main programs: the Coffee Stabilization Fund, which was a price floor program aimed at guaranteeing a price of 340 dollars per ton of cherry coffee to help farmers to cover the production costs of plantations located in areas suitable to produce high quality coffee (usually areas located at elevations higher than 900 meters above sea level); and the Productive Reconversion Program which was focused on partially funding the removal of coffee plantations in low-altitude areas that cannot produce high quality coffee. Coffee growers in the study area were potential beneficiaries of the Productive Reconversion Program, and were eligible to receive two lump-sum transfers of around 80 dollars during two years to transition to a different land use. However participation was generally low because that amount is not enough to pay the costs associated with the removal of coffee plantations, let alone the expenses related to establish a new crop. Additionally, the information about the participants in that program is confidential, so we do not have data to explicitly model the effects of that policy in the observed land use decisions. Therefore the potential effects of such policy in the observed landscape dynamics are captured in the unobservable components.

Both of these programs fundamentally were income support programs for small holder farmers although they have two different objectives. A natural question that arises is: how well the Coffee Stabilization Fund would perform as conservation tools focused on preservation of shade canopy in low land areas? A study implemented by Avalos-Sartorio and Blackman (2010) to investigate the performance of the Coffee Stabilization Fund as a conservation tool in the higher elevations, finds that farmers in areas that were more likely to experience coffee plantation removal had low levels of participation in the program, and that the price floor was not high enough to affect land use change decisions. Given the observed tree canopy loss in the present analysis, and as an additional mechanism to contrast the performance of the static and dynamic models, here I evaluate the welfare implications and landscape configurations resulting from implementation of a similar price floor policy in my study area. Specifically I evaluate a policy that would have extended the Coffee Stabilization Fund price floor to the lowland regions throughout the period of analysis. Figure 5.6. shows the fallow-normalized agroforestry price index used in the dynamic model (which is entirely based on the average price per ton received by coffee growers), along with its average value. The same figure also shows the inter-temporal trend of a normalized price floor corresponding to a price of 340 dollars per ton of coffee. The upward trend in the normalized price floor is generated by the continuous decline in the price index of the FA category observed during the study period. Because the land use profitability estimation in the mixed conditional–multinomial logit and standard multinomial logit models rely on assumptions that differ from those in the dynamic model, in this section I use the parameter estimates for the dynamic model with discount factors of 5% and 57% to implement an unbiased comparison of the effects of the simulated price floor policy, assuming that the dynamic model with a discount factor of 5% is a close representation of a static model.

The analysis of the welfare effects and land use decisions resulting from the aforementioned price floor policy was implemented using two procedures based on the analytical expression of the expected payoff derived from discrete choice logit models

$$E\left[\max\left\{v_1+\varepsilon_1,\ldots,v_J+\varepsilon_J\right\}\right]=\ln\sum_{j=1}^{J}\exp\left(v_j\right)+0.57722$$

where v_j is described in equation 5, and J indicates the total number of land use categories contained in the choice set. The first procedure consisted on using the observed realizations of the AG price index and the parameter estimates generated with the dynamic model with discount factors of 5% and 57%, to estimate the baseline welfare and land allocation without the policy. After computing the baseline the observed AG prices located below the normalized floor price were replaced by the corresponding floor value. Additionally, under the assumption that farmers would know that the price floor would be implemented during all the study period, the structure of the AR1 process used to estimate the transition probability matrix of the AG price was re-estimated to account for the change in the minimum expected price and long run mean of the AR1 process. After those adjustments, the welfare effects and land use decisions generated under the simulated policy were computed using the two reference discount factors. The welfare estimates are shown in figure 5.7, and the proportion of land allocated to each land use category are presented in figure 5.8.

Figure 5.6 Time series trends of the normalized Agroforestry price and the normalized price floor.

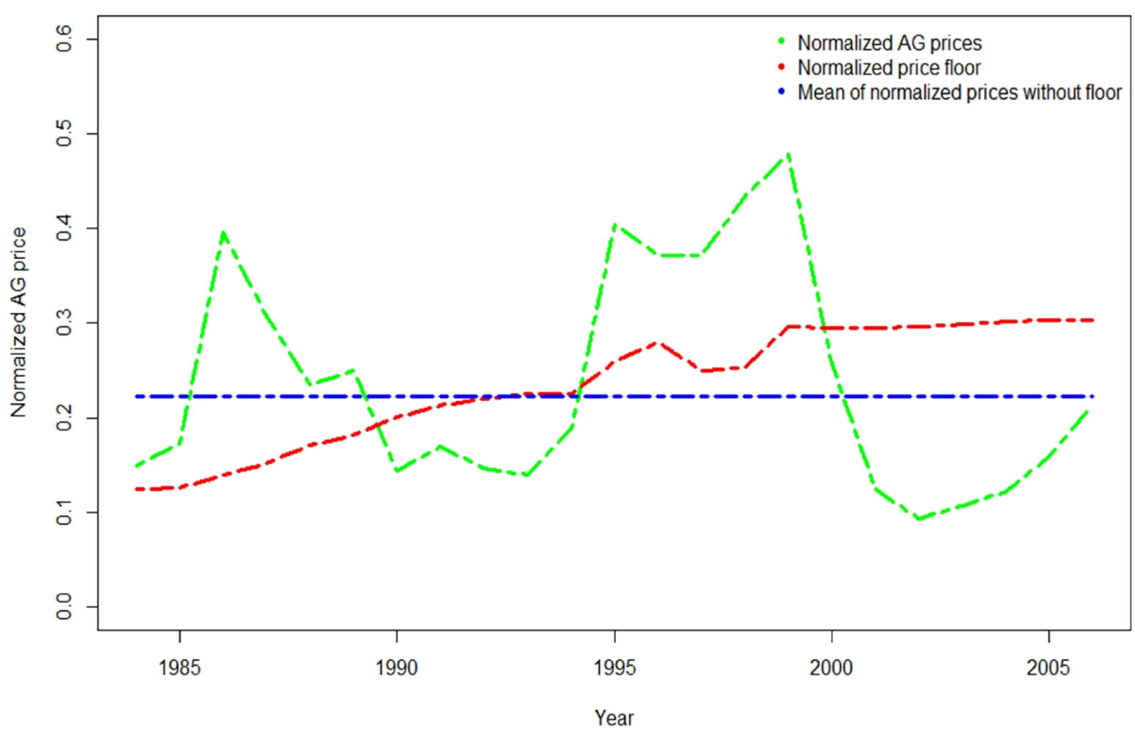


Figure 5.7. Welfare change of a simulated price floor policy using the observed price realizations as reference for the analysis.

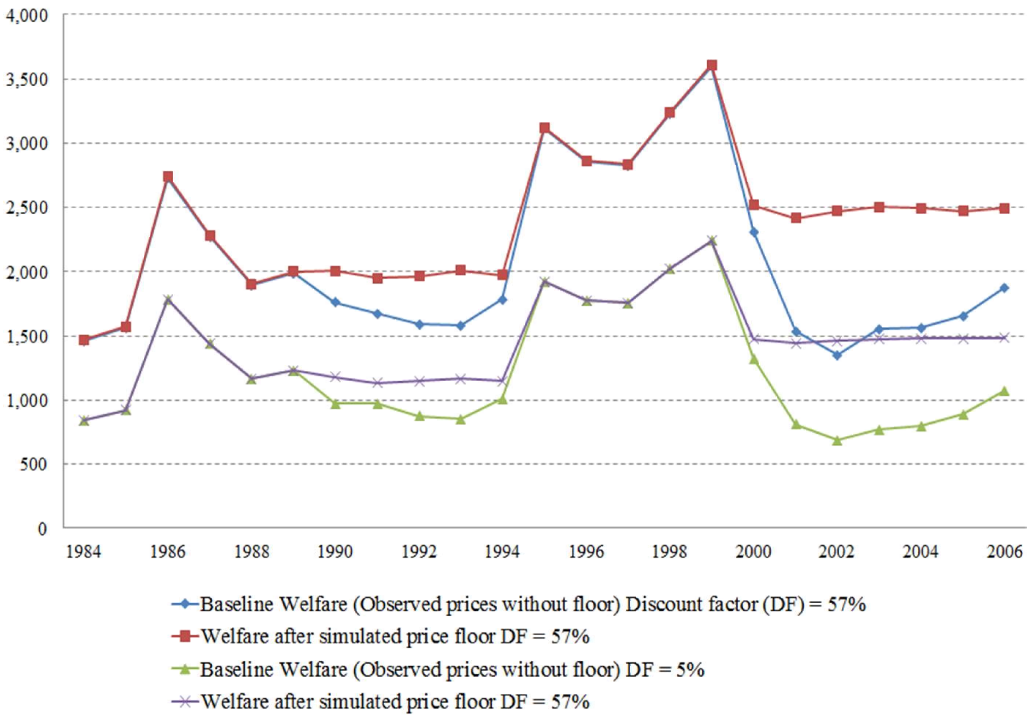
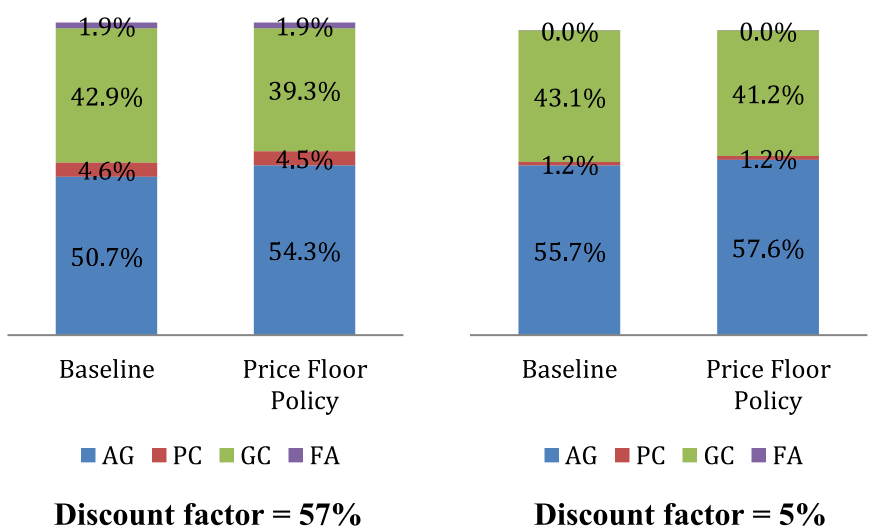


Figure 5.8. Aggregated land use proportions with and without the simulated price floor policy



The results generated with this procedure indicate that, as expected, the simulated price floor policy would produce higher welfare levels than the baseline, especially during periods in which the market price plummeted below the simulated price floor. On the other hand, the welfare estimates of the dynamic model that assumes a 57% discount factor are 39.8% higher than the estimations generated assuming a discount factor of 5%. With regard to the impact of the simulated policy as a conservation tool to reduce the removal of shadow coffee plantations, the model at a discount rate of 57% indicates an increase in the agroforestry category of 3.6%, while the model with a 5% discount factor indicates an increase of 2% in the land corresponding to the same land use category.

Given that the aforementioned procedure used to analyze the impact of the simulated price floor policy is based only on the observed realization of the agroforestry price index, an additional analysis was implemented under the same general assumptions but this time simulating different price paths and estimating the average baseline welfare level and land use decisions, as well as identifying the corresponding average results of the simulated price floor policy. In general terms this second procedure followed the next steps:

1. Compute 500 random price paths based on the AR1 structure of the normalized AG price index and store those values as representatives of prices without policy.

2. In each price path replace the values that are below the normalized price floor with the price floor value and store those price paths to evaluate the impacts of the simulated price policy.

3. Re-estimate the structure of the AR1 process used to estimate the transition probability matrix of the AG price, to account for the change in the minimum expected price and long run mean of the AR1 process.

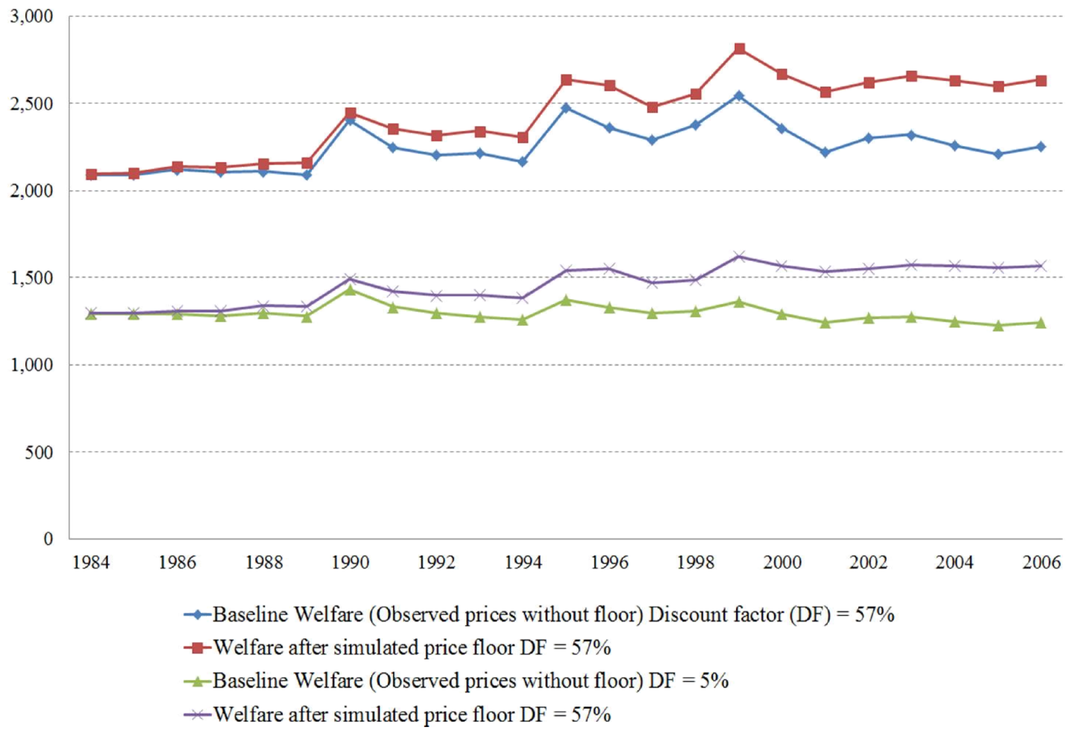
4. Estimate the value function corresponding to the discretized state space and the modified AGP price transition probability matrix.

5. Use cubic spline interpolation to estimate the optimal decision rule at each of the simulated random AG prices as well as welfare for each sample parcel, and average the results.

Figure 5.9. shows the average welfare levels, and figure 5.10 shows the average land use proportions estimated with the simulated random price paths. We can observe that on average, at both levels of the discount factor, the model predicts an increase in welfare derived from the simulated price floor. On the other hand, the dynamic model with a 57% discount factor predicts welfare estimates that are 40.06% higher than the estimations generated with a discount factor of 5%, which is very close to the 39.8% detected with the first procedure. Nevertheless, a significant difference with the previous procedure is that the results from the low discount factor model indicate that the price policy would not alter the proportion of land allocated to agroforestry production. On the other hand, the analysis with the model that uses a 57% discount factor indicates that an

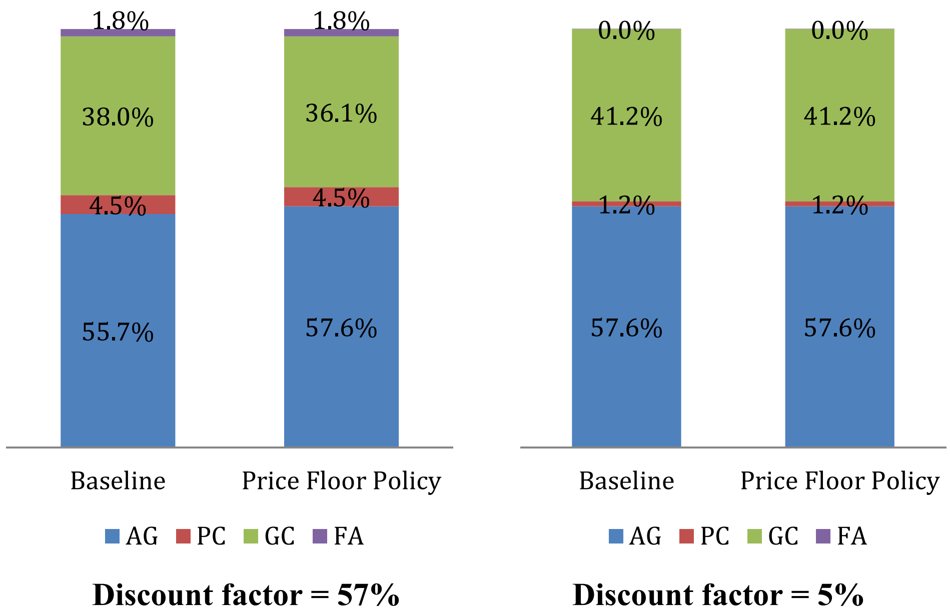
increase of around 2% in the agroforestry category could have been observed in the study area if a price floor policy similar to the Coffee Stabilization Fund had been implemented during all the period of analysis. Furthermore, this model indicates that the additional percentage of coffee plantations correspond to land uses that should have been allocated to the GC category; a land use type that requires very low tree canopy density.

Figure 5.9. Average welfare change of a simulated price floor policy using 500 simulated price paths.



Overall the results appear to support the reports by Avalos-Sartorio and Blackman (2010) that indicate that the price floor policy promoted by the Mexican government did not target coffee growing areas that were more affected by the international coffee crisis. Although their analysis focuses on high altitude coffee growing regions, and during the period in which the policy was actually implemented, the simulation exercise implemented in this section indicates that such a policy would have helped to preserve agroforestry plantations in marginal coffee growing areas if agent's discount factor is in fact around 57%.

Figure 5.10. Aggregated land use proportions with and without the simulated price floor policy using 500 simulated random price paths



5.5 Summary

This chapter presents a description of the implementation of a discrete choice dynamic programming model of land use decisions. The empirical application focuses on investigating the structure of the decision making process followed by farmers in a Mexican coffee growing region during the period 1984 – 2006.The estimation of the maximum likelihood parameter estimates associated with the behavioral process under analysis is implemented using Rust's (1987) nested fixed point algorithm. For comparison purposes I implement a mixed conditional multinomial logit model and a multinomial logit approach using the same reduced set of independent variables considered in the dynamic analysis.

The results indicate that the dynamic model can provide highly accurate predictions for the data used to calibrate the model at low levels of the discount factor. Nevertheless, the dynamic model at a discount factor of 57% provides a substantially better fit to the observed inter-temporal trajectory of land uses and provides greater consistency with the presumed structure of land use decisions. The directions of the marginal effects corresponding to parcel specific characteristics that are considered constant during the period of analysis have in general the expected directions in both the static and dynamic models. The analysis also indicates that the implemented static models fail to produce theoretically consistent results for the revenue variable corresponding to the PC category. The dynamic approach even at a 5% discount factor has the ability to produce marginal effects for the PC revenue variable that have the expected direction. Nevertheless, none of the tested models can provide theoretically consistent marginal effects for the revenue variable of the GC category. In the particular framework defined by the observed inter-temporal land use trajectories and average prices received by farmers in the study region during the period of analysis, the empirical application of the dynamic model is helpful for understanding land use decisions during the relatively long period of sustained low output prices. Such situation was observed to produce counter-intuitive parameter estimates in the implemented static models. Furthermore, analysis of a simulated price floor policy aimed at reducing deforestation in agroforestry systems indicates that welfare levels estimated with a dynamic model are around 40% higher than under a model that assumes a discount factor close to zero, and predict a slight increase in the proportion of agroforestry parcels. Nevertheless, further work needs to be done to improve the dynamic analysis; for instance, by relaxing the risk neutrality assumption implicit in the modeling approach used in this chapter, or by evaluating alternative subjective price expectation procedures to the AR1 structure used in the empirical model.

Chapter 6

Conclusions

Through intrinsically dynamic and multidirectional patterns and processes, socioeconomic and environmental variables interact to influence agents' land use decisions. A comprehensive analysis of the complex interlinkages of the driving forces of land use change observed at different temporal and spatial scales, and across systems, can be difficult to model. One commonly used approach to handle this issue is to divide the relevant patterns and processes into modules that are separately analyzed by economists, ecologists, etc. Once the relevant interactions are represented, at the module-level, the results can be aggregated and linked to produce an accurate representation of the landscape configuration process. In this thesis the attention centers on addressing methodological issues that impact the performance of spatially explicit discrete choice agent-based land use models that are estimated with remotely sensed data. On the empirical side this research focuses on the study of land use transitions between coffee-based agroforestry systems, perennial crops (citrus and banana), grasslands and corn, and fallow lands in a Mexican region in which relatively high rates of tree canopy removal were observed as a result of the clearing of shade-grown coffee plantations.

As a starting point of the analyses implemented in this thesis, a Mixed Conditional – Multinomial Logit (MCML) model was implemented to highlight assumptions and limitations associated with the use of spatially explicit discrete choice random utility models in the analysis of land use decisions. Particularly, the structures that were used to model agents' price and yield expectations are shown to be limited approximations of two key elements used in the estimation of the profitability of land use choices. The estimated marginal effects corresponding to the parcel specific land use drivers have in general the expected directions. However, the model produces theoretically inconsistent parameter estimates for the revenue variable associated with three out of four land uses considered in the analysis. Particularly this model cannot produce parameter estimates that could be used to provide an economically rational explanation for the observed increasing proportion of perennial crops in the study region, and the concomitant decrease in the associated market prices. On the other hand, the MCML model can accurately predict 59% of the observed land use decisions at the parcel level, which is similar to results obtained with a multinomial logit approach that uses normalized prices and considers that revenue can be modeled as a parcel specific variable (Baerenklau et al., 2012). As a reference to compare the predictive performance of the MMCL model, random guessing should produce 25% accurate predictions, given that the choice set is composed of four land use categories. Nevertheless, an analysis of the ability of the model to predict decisions at each of the observation points for which there is land use information indicates that the model accuracy decreases through time, a trend that is observed to continue in an out-of-sample validation exercise using land use data collected in 2011.

To investigate whether the counterintuitive marginal effect estimates for most of the revenue variables are generated from misclassified land use data, a Latent

Multinomial Logit model was implemented. This approach allowed the identification of land use observations that have a high likelihood of being wrongly classified. Furthermore, it provided evidence that the procedure implemented to construct a fallow lands category is likely misinterpreting temporary increases in biomass in the grasslands and corn category as land use change. Such changes in biomass instead may be the result of a production system that requires temporary land abandonment as a mechanism to recover soil productivity, or simply an indication that the land had not been maintained during the time in which the remotely sensed data were collected. A reconfiguration in the dataset based on the results from this model indicates that the original sample over-represents the proportion of land classified in the agroforestry category, under represents the percentage of land classified as grass and corn, and over represents the observations classified as fallow lands. With regard to the parameter estimates, the results indicate that a reclassification of the land use data, based on the latent multinomial logit model, increases the magnitudes of the marginal effects of the analyzed land use drivers in the theoretically expected directions. Particularly, the marginal effect of changes in revenue associated with the fallow lands category becomes statistically significant and shows the theoretically expected direction. Nevertheless, this model still fails to produce a theoretically consistent explanation of the observed decisions to increase the proportion of perennial crops in the study area despite the continuous drop in price of the crops associated with that land use. Unfortunately, the dataset cannot be used to evaluate the performance of the model in the estimation of misclassified data. Simulated land use observations or a highly accurate (i.e., groundtruthed) empirical land use dataset could be used to artificially introduce misclassification in the reference data and then test the ability of the LMNL approach to reconstruct the original land use classification.

A drawback of the implemented LMNL model is that it does not explicitly consider how land use classifications at the pixel level are temporally linked, i.e., it does not incorporate directly in the analysis the state of each pixel before and after the observation year to help the model identify more accurate states of the pixel by removing classifications that do not present economically rational sequences. In this context, the misclassification problem could be analyzed with a state-space model, and the Kalman filter could be used to keep track of the state of each pixel at every observation point (i.e., the land use assigned to each pixel at any point in time), to produce information about the uncertainty of the land use classification assigned to each pixel, and generate parameter estimates (Knapp & Konyar, 1991). This would be a logical extension of the work accomplished in this thesis.

Since static discrete choice models require limiting assumptions that potentially oversimplify the behavioral process followed by landowners, effectively undermining the role of future expectations in current decision making, this thesis considers that the interactions between immediate and underlying driving forces of landscape change are better explained using a stochastic-dynamic (i.e., forward-looking) framework. This justifies the implementation of a modeling approach focused on identifying the structure of a more realistic behavioral model that assumes that land managers are forward-looking and act to maximize their discounted flow of current and future expected utility within a stochastic environment.

To reduce the dimensionality of the model, a reduced set of explanatory variables was used in the estimation of the dynamic model. Elevation, distance from a parcel to the nearest market, and welfare status were used as parcel specific characteristics that are relatively constant during the study period. These land use drivers were identified during the calibration process as relevant variables to the analysis. Furthermore, transitions rules were determined to control for the evolution of the age of the plantations, and the price elicitation process followed by land managers was assumed to have an AR structure approximated by a Markov chain.

The structural parameters that govern the decision process followed by farmers to allocate their land between competing options were estimated using the nested-fixed point algorithm. One of the key elements in that modeling approach is the assumed discount factor that agents use to estimate the net present value of expected flows of payoffs corresponding to the different land use paths that could be implemented in their parcels. To identify the discount factor level that best approximates the "true" discount factor used by agents in the study region, a grid search procedure was implemented. Typically the discount factor that maximizes the nested likelihood function or that produces better predictions is selected. Nevertheless, the results indicate that as the discount factor approaches zero the log-likelihood value and the percentage of per parcel accurate predictions in the data used to calibrate the model both increase. Additionally, the dynamic model with a discount factor within the interval $(0, 0.75]$ is found to produce highly accurate out-of sample predictions. Further analysis into the effects of different discount factor levels in the estimated land use proportions revealed that a discount factor of 57% provides the closest approximation to the observed inter-temporal landscape dynamics. Despite low values of the discount factor generating highly accurate predictions at the parcel level and higher values for the log-likelihood function, those types of measure of fit should not be used to assess whether an underlying decision process is dynamic (Provencher & Baerenklau, 2005). Since land use decisions are inherently dynamic the discount factor of 57% is considered to be more consistent with forward-looking behavior. Furthermore that discount factor was found to perform relatively well under the commonly used aforementioned selection mechanisms.

The comparison between static and dynamic models shows that the directions of the marginal effects corresponding to time-invariant parcel-specific variables generally have the expected directions independent of the selected modeling approach. More importantly, the marginal effect estimates for the cash crop revenue variables (i.e., agroforestry and perennial crops) have the expected direction in the dynamic model. By contrast the myopic modeling approaches generate counter-intuitive results for the PC revenue variable, which make those results unusable for policy design.

A policy simulation exercise showed the sensitivity of welfare estimates to the discount factor selected as representative of the true value used by decision makers. The results of the simulated price floor policy with discount factors of 57%, and 5% indicate that the price floor set by the Mexican government after 2001 to help coffee growers maintain their plantations seems to be insufficient to influence agents' decisions in the study region, which is representative of the areas that were more severely affected by the international coffee crisis. These results are consistent with the analysis implemented in a different Mexican coffee growing region by Avalos-Sartorio and Blackman (Ávalos-Sartorio & Blackman, 2010).

A more robust analysis of agent based land use decisions in the study area could be achieved in several ways. First, the latent model specification could be incorporated in the discrete choice dynamic model. Such an implementation has the potential to provide a more realistic and accurate analysis of the underlying structure of the behavioral process followed by land managers, when the information available to the analyst is suspected to have a significant percentage of misclassified data. Unfortunately, the computational cost of a latent discrete choice multinomial logit model would be high and thus is not attempted here. Second, controlling explicitly for spatial autocorrelation in the modeling approach, instead of implementing systematic random sampling to produce a sample of spatially independent observations, could help to improve the robustness of the modeling assumptions and the validity of the estimated marginal effects. Third, during recent years the approach of using probability formats to elicit expectations has become a common approach in the economic arena, and several studies offer empirical evidence that supports its validity (Delavande, Giné, & McKenzie, 2011; Manski, 2004). Therefore, to provide a more realistic assessment of the process followed by decision makers in the estimation of future revenue expectations, survey methods could be implemented to elicit the subjective price expectation process followed by representative agents in the region under analysis. Of course, ad hoc rules (e.g. moving averages, ARIMA processes) can be used as a proxy of the expectation formation process, but these rules might not be empirically justified. Alternatively, the AR1 parameters could be endogenously estimated using the nested fixed-point algorithm[6]. Fourth, a modeling approach based on recursive preferences could help to relax the risk neutrality assumption implicit in discrete choice dynamic models, which is difficult to justify in land uses change analysis with categories that involve high up front investments and volatile prices, or that are implemented in relatively poor regions. In general, future research derived from this thesis should control for an adequate definition of the environment that surrounds the agents (e.g. land use distribution, soil quality, infrastructure, governmental policies, climate patterns, water availability, etc.); the state of the agents in that environment (e.g. property regimes, institutions, social connections), and their decision-making criteria (e.g., profit maximizing agents with rational or subjective expectations, degree of risk aversion, influence of non-monetary elements in their payoff function, etc.).

[6] This approach was tested during model calibration stages but it was observed that gradient based constrained optimization solvers had difficulty moving away from the starting values of the AR1 parameter estimates.

References:

Adiku, S. G. K., Kumaga, F. K., Tonyigah, A., & Jones, J. W. (2009). Effects of crop rotation and fallow residue management on maize growth, yield and soil carbon in a savannah-forest transition zone of Ghana. *The Journal of Agricultural Science*. doi:10.1017/S002185960900851X

Albers, H. J., Avalos-Sartorio, B., Batz, M. B., & Blackman, A. (2006). Maintenance Costs, Price Uncertainty, and Abandonment in Shade-Grown Coffee Production: Coastal Oaxaca, Mexico. In *Environmental and Resource Economists 3rd World Congress* (p. 39). Retrieved from http://www.webmeets.com/ERE/WC3/Prog/

Andersen, L. E. (1996). The Causes of Deforestation in the Brazilian Amazon. (R. D. Ed, Ed.)*The Journal of Environment Development*, *5*(3), 309–328. doi:10.1177/107049659600500304

Angelsen, A. (2007). Forest cover change in space and time : combining the von Thünen and Forest Transition. *World Bank Policy Research Working Paper*, (4117), 1–43.

Angelsen, A., & Kaimowitz, D. (1999). Rethinking the causes of deforestation: lessons from economic models. *The World Bank Research Observer*, *14*(1), 73–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12322119

Aoki, K., & Suvedi, M. (2012). Coffee as a Livelihood Support for Small Farmers: A Case Study of Hamsapur Village in Nepal. *Journal of International Agricultural and Extension Education*, *19*, 16–29. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84870033895&partnerID=40&md5=549570985480850ce41bbb87a8521b3c

Arslan, A., & Taylor, J. E. (2008). Farmers ' Subjective Valuation of Subsistence Crops : The Case of Traditional Maize in Mexico by Farmers ' Subjective Valuation of Subsistence Crops : The Case of Traditional Maize in Mexico. Retrieved from http://purl.umn.edu/44488

Ávalos-Sartorio, B., & Blackman, A. (2010). Agroforestry price supports as a conservation tool: Mexican shade coffee. *Agroforestry Systems*, *78*(2), 169–183. doi:10.1007/s10457-009-9248-4

Baerenklau, K., Ellis, E. A., & Marcos-Martínez, R. (2012). Economics of land use dynamics in two Mexican coffee agroforests: implications for the environment and inequality. *Investigación Económica*, *LXXI*(279), 93–124.

Baerenklau, K., & Provencher, B. (2005). Static modeling of dynamic recreation behavior: Implications for prediction and welfare estimation. *Journal of Environmental Economics and Management*, *50*(3), 617–636. doi:10.1016/j.jeem.2005.03.001

Bakker, M. M., Govers, G., Kosmas, C., Vanacker, V., Oost, K. van, & Rounsevell, M. (2005). Soil erosion as a driver of land-use change. *Agriculture, Ecosystems & Environment*. doi:10.1016/j.agee.2004.07.009

Begueria, S. (2006). Changes in land cover and shallow landslide activity: A case study in the Spanish Pyrenees. *Geomorphology*, *74*(1-4), 196–206. doi:10.1016/j.geomorph.2005.07.018

Bhagwat, S. A., Willis, K. J., Birks, H. J. B., & Whittaker, R. J. (2008). Agroforestry: a refuge for tropical biodiversity? *Trends in ecology & evolution (Personal edition)*, *23*(5), 261–267.

Blackman, A., Albers, H. J., Avalos-Sartorio, B., & Murphy, L. C. (2008). Land cover in a managed forest ecosystem: Mexican shade coffee. *American Journal of Agricultural Economics*, *90*(1), 216–231. doi:10.1111/j.1467-8276.2007.01060.x

Blackman, A., Ávalos-Sartorio, B., & Chow, J. (2012). Land cover change in agroforestry: Shade coffee in El salvador. *Land Economics*, *88*(1), 75–101. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84855582595&partnerID=40&md5=5878b1594f6c4ddd816c0292b6344ec9

Bonan, G. B., DeFries, R. S., Coe, M. T., & Ojima, D. S. (2004). Land use and climate. In M. A. Gutman, G.; Janetos, A.C.; Justice, C.O.; Moran, E.F.; Mustard, J.F.; Rindfuss, R.R.; Skole, D.; Turner II, B.L.; Cochrane (Ed.), *Land Change Science. Observing, Monitoring and Understanding Trajectories of Change on the Earth's Surface.* (pp. 301–314). Netherlands.: Springer.

Brady, M., & Irwin, E. (2011). Accounting for Spatial Effects in Economic Models of Land Use: Recent Developments and Challenges Ahead. *Environmental and Resource Economics*, *48*(3), 487–509. doi:10.1007/s10640-010-9446-6

Brooks, J. (2010). Development , Poverty Reduction Agricultural Policy Choices in Developing Countries : A Synthesis. Paris. Retrieved from http://www.oecd.org/tad/agricultural-policies/46340461.pdf

Bürgi, M., Hersperger, A. M., & Schneeberger, N. (2004). Driving forces of landscape change – current and new directions. *Landscape Ecology*, *19*(8), 857–868. doi:10.1007/s10980-004-0245-8

Card, D. E. (2008). Dynamic discrete choice. Retrieved from http://emlab.berkeley.edu/users/webfac/card/econ250c_f08/lecture8.pdf

Caudill, S. B. (2006). A logit model with missing information illustrated by testing for hidden unemployment in transition economies. *Oxford Bulletin of Economics and Statistics*, *68*(5), 665–677.

Caudill, S. B., Ayuso, M., & Guillen, M. (2005). Fraud Detection Using a Multinomial Logit Model With Missing Information. *Journal of Risk and Insurance*, *72*, 539–550. doi:10.1111/j.1539-6975.2005.00137.x

Caudill, S. B., Groothuis, P. A., & Whitehead, J. C. (2011). The Development and Estimation of a Latent Choice Multinomial Logit Model with Application to Contingent Valuation. *American Journal of Agricultural Economics*, *93*(4), 983–992. doi:10.1093/ajae/aar030

Caudill, S. B., & Mixon Jr., F. G. (2005). Analysing misleading discrete responses: A logit model based on misclassified data. *Oxford Bulletin of Economics and Statistics*, *67*(1), 0305–9049 ST – Practitioners' Corner: Analysing M.

Chardin, A., & Perez, P. Unsupervised image classification with a hierarchical EM algorithm. , 2 Proceedings of the Seventh IEEE International Conference on

Computer Vision 969 (1999). Published by the IEEE Computer Society. doi:10.1109/ICCV.1999.790353
Chomitz, K. M., & Gray, D. A. (1996). Roads, land use, and deforestation: a spatial model applied to Belize. *The World Bank Economic Review*, *10*, 487–512. doi:10.1093/wber/10.3.487
CONAPO. (1998). *Índices de marginación, 1995* (1st ed.). Mexico City: Consejo Nacional de Población.
CONAPO. (2006). *Índices de Marginación 2005* (1st ed.). Mexico City: Consejo Nacional de Población. Retrieved from http://www.conapo.gob.mx/es/CONAPO/Indices_de_marginacion_2005_
CONAPO. (2011). *Indice de Marginacion por Entidad Federativa y Municipio, 2010.* Mexico City: Consejo Nacional de Población.
Conservation International. (2013). *Mexico: Deforestation guide.* (p. 2). Retrieved from http://www.conservation.org/how/science/Documents/DeforestationGuide_CommodititySourcing_Mexico.pdf
Cornwall, A., & Jewkes, R. (1995). What is participatory research? *Social science & medicine*, *41*(12), 1667–1676.
Croissant, Y. (2008). Estimation of Multinomial Logit Models in R : The mlogit Packages. *Data Management*. Retrieved from http://cran.r-project.org/web/packages/mlogit/vignettes/mlogit.pdf
De Janvry, A., Fafchamps, M., & Sadoulet, E. (1991). Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. *The Economic Journal*, *101*(409), 1400–1417. doi:10.2307/2234892
De Pinto, A., & Nelson, G. C. (2008). Land Use Change with Spatially Explicit Data: A Dynamic Approach. *Environmental and Resource Economics*, *43*(2), 209–229. doi:10.1007/s10640-008-9232-x
Delavande, A., Giné, X., & McKenzie, D. (2011). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics*, *94*, 151–163. doi:10.1016/j.jdeveco.2010.01.008
Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B Methodological*, *39*(1), 1–38. doi:10.2307/2984875
Dinata Putra, A. E., Verbist, B., & Budidarsono, S. (2005). Factors driving land use change: Effects on watershed functions in a coffee agroforestry system in Lampung, Sumatra. *Agricultural Systems*. doi:10.1016/j.agsy.2005.06.010
Du, P., Xia, J., Zhang, W., Tan, K., Liu, Y., & Liu, S. (2012). Multiple Classifier System for Remote Sensing Image Classification: A Review. *Sensors (Peterboroug*, *13*, 4764–4792. doi:10.3390/s120404764
Dunn, R., & Harrison, A. R. (1993). Two-Dimensional Systematic Sampling of Land Use. *Journal of the Royal Statistical Society Series C Applied Statistics*, *42*(4), 585–601. Retrieved from http://www.jstor.org/stable/2986177
Ellis, E. A., Baerenklau, K. A., Marcos-Martínez, R., & Chávez, E. (2010). Land use/land cover change dynamics and drivers in a low-grade marginal coffee growing region

of Veracruz, Mexico. *Agroforestry Systems*, *80*(1), 61–84. doi:10.1007/s10457-010-9339-2

Enders, C. K. (2010). *Applied missing data analysis*. *Methodology in the social sciences* (p. 377). Guilford Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=MN8ruJd2tvgC&oi=fnd&pg=PA1&dq=Applied+missing+data+analysis&ots=dHhyqZXcx2&sig=xJcvimM2PQvNu4craFvBVemuLrg

Escamilla Prado, E. (2007). *Influencia de los factores ambientales, genéticos. agronómicos y sociales en la calidad del café orgánico en México.* Colegio de Postgraduados. Retrieved from http://www.biblio.colpos.mx:8080/jspui/handle/10521/1625

FAO. (2010). Global Forest Resources Assessment 2010. Main Report. *FAO FORESTRY PAPER*. Rome: Food and Agriculture Organization of the United Nations. doi:ISBN 978-92-5-106654-6

Fu, B.-J., Zhang, Q.-J., Chen, L.-D., Zhao, W.-W., Gulinck, H., Liu, G.-B., … Zhu, Y.-G. (2006). Temporal change in land use and its relationship to slope degree and soil type in a small catchment on the Loess Plateau of China. *CATENA*. doi:10.1016/j.catena.2005.07.005

Gay, C., Estrada, F., Conde, C., Eakin, H., & Villers, L. (2006). Potential Impacts of Climate Change on Agriculture: A Case of Study of Coffee Production in Veracruz, Mexico. *Climatic Change*, *79*(3-4), 259–288. doi:10.1007/s10584-006-9066-x

Geist, H. J., & Lambin, E. F. (2002). Proximate Causes and Underlying Driving Forces of Tropical Deforestation. *BioScience*, *52*(2), 143–150. doi:10.1641/0006-3568(2002)052[0143:PCAUDF]2.0.CO;2

Griffith, D. A. (2009). Spatial Autocorrelation. In *The International Encyclopedia of Human geography*. Elsevier Inc. Retrieved from http://www.elsevierdirect.com/brochures/hugy/SampleContent/Spatial-Autocorrelation.pdf

Guo, Q., Ma, K., Yang, L., & He, K. (2010). Testing a Dynamic Complex Hypothesis in the Analysis of Land Use Impact on Lake Water Quality. *Water Resources Management*, *24*(7), 1313–1332. doi:10.1007/s11269-009-9498-y

Haines-Young, R. (2009). Land use and biodiversity relationships. *Land Use Policy*, *26*(S1), S178–S186. doi:10.1016/j.landusepol.2009.08.009

Hansen, A. J., DeFries, R. S., & Turner, W. (2004). Land Use Change and Biodiversity: A Synthesis of Rates and Consequences during the Period of Satellite Imagery. In *Land Change Science* (Vol. 6, pp. 277–299). Springer Netherlands. doi:10.1007/978-1-4020-2562-4

Hicks, R. L., & Schnier, K. E. (2006). Dynamic random utility modeling: A Monte Carlo analysis. *American Journal of Agricultural Economics*, *88*(4), 816–835.

Hongzan, J., Yanfei, Z., Liangpei, Z., & Pingxiang, L. (2011). Unsupervised remote sensing image classification using an artificial DNA computing. *2011 Seventh International Conference on Natural Computation*. IEEE. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6022338

Huang, W., Luukkanen, O., Johanson, S., Kaarakka, V., Räisänen, S., & Vihemäki, H. (2002). Agroforestry for biodiversity conservation of nature reserves: functional group identification and analysis. *Agroforestry Systems*, *55*(1), 65–72. doi:10.1023/A:1020284225155

INEGI. (1998). Curvas de nivel para la República Mexicana. Mexico City: Instituto Nacional de Estadística, Geografía e Informática.

INEGI. (1999). *Cartas vectoriales 1:20 000*. Mexico City.

INEGI. (2001). *XII Censo General de Población y Vivienda 2000*. Mexico City.

INEGI. (2009). *Estadísticas históricas de México*. Retrieved from http://www.inegi.org.mx/prod_serv/contenidos/espanol/bvinegi/productos/integracion/pais/historicas10/Tema6_Salarios.pdf

Jiang, L., Bergen, K. M., Brown, D. G., Zhao, T., Tian, Q., & Qi, S. (2008). Land-cover Change and Vulnerability to Flooding near Poyang Lake , Jiangxi Province , China. *Photogrammetric Engineering and Remote Sensing*, *74*(6), 775–786.

Jordan-Garcia, A., Collazo, J. A., Borkhataria, R., & Groom, M. J. (2012). Shade-grown coffee in Puerto Rico: Opportunities to preserve biodiversity while reinvigorating a struggling agricultural commodity. *Agriculture, Ecosystems & Environment*. doi:10.1016/j.agee.2010.12.023

Keane, M. P., & Wolpin, K. I. (2009). Empirical applications of discrete choice dynamic programming models. *Review of Economic Dynamics*, *12*(1), 1–22. doi:10.1016/j.red.2008.07.001

Knapp, K. C., & Konyar, K. (1991). Perennial Crop Supply Response : A Kalman Filter Approach. *American Journal of Agricultural Economics*, *73*(3), 841–849.

Kolawole, G. O., Salako, F. K., Idinoba, P., Kang, B. T., & Tian, G. (2005). Long-term effects of fallow systems and lengths on crop production and soil fertility maintenance in West Africa. *Nutrient Cycling in Agroecosystems*. doi:10.1007/s10705-004-1927-y

Kuan, C.-M., Leisch, F., & Hornik, K. (2000). Monitoring Structural Changes with the Generalized Fluctuation Test. *Econometric Theory*, *16*(6), 835–854. doi:10.1017/S0266466600166022

Kursten, E. (2000). Fuelwood production in agroforestry systems for sustainable land use and CO2-mitigation. *Ecological Engineering*, *16*, S69–S72.

Lal, R. (2001). Soil degradation by erosion. *Land Degradation & Development*, *12*(6), 519–539. doi:10.1002/ldr.472

Lal, R. (2004). Soil carbon sequestration to mitigate climate change. *Geoderma*, *123*(1-2), 1–22. doi:10.1016/j.geoderma.2004.01.032

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Folke, C., … Veldkamp, T. A. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, *11*(4), 261–269. doi:10.1016/S0959-3780(01)00007-3

Lewis, J., & Runsten, D. (2008). Is Fair Trade-Organic Coffee Sustainable in the Face of Migration? Evidence from a Oaxacan Community. *Globalizations*. doi:10.1080/14747730802057738

Maestas, J. D., Knight, R. L., & Gilgert, W. C. (2003). Biodiversity across a Rural Land-Use Gradient. *Conservation Biology*, *17*(5), 1425–1434. doi:10.1046/j.1523-1739.2003.02371.x

Manski, C. F. (2004). Measuring Expectations. *Econometrica*, *72*, 1329–1376. doi:10.1111/j.1468-0262.2004.00537.x

Mas, J.-F., Velázquez, A., Díaz-Gallegos, J. R., Mayorga-Saucedo, R., Alcántara, C., Bocco, G., … Pérez-Vega, A. (2004). Assessing land use/cover changes: a nationwide multidate spatial database for Mexico. *International Journal of Applied Earth Observation and Geoinformation*, *5*(4), 249–261. doi:10.1016/j.jag.2004.06.002

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. (2007). Agent-based land-use models: a review of applications. *Landscape Ecology*, *22*(10), 1447–1459. doi:10.1007/s10980-007-9135-1

McLachlan, G. J., & Krishnan, T. (1997). *The EM Algorithm and Extensions*. (J. W. Sons, Ed.)*New York* (Vol. 274, p. 359). Wiley. doi:10.1002/9780470191613

Melgani, F. Contextual reconstruction of cloud-contaminated multitemporal multispectral images. , 44 IEEE Transactions on Geoscience and Remote Sensing 442–455 (2006). doi:10.1109/TGRS.2005.861929

Messer, K. D., Kotchen, M. J., & Moore, M. R. (2000). Can shade-grown coffee help conserve tropical biodiversity? A market perspective. *Endangered Species Update*, *17*(6), 125–131.

Meyer, W. B., & Turner B L, I. I. (1992). Human Population Growth and Global Land-Use/Cover Change. *Annual Review of Ecology and Systematics*, *23*(1), 39–61. Retrieved from http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.es.23.110192.000351

Millington, J., Romero-calcerrada, R., Wainwright, J., Perry, G., Societies, A., & Simulation, S. (2008). An Agent-Based Model of Mediterranean Agricultural Land-Use/Cover Change for Examining Wildfire Risk. *Journal of Artificial Societies and Social Simulation*, *11*(4), 4. Retrieved from http://jasss.soc.surrey.ac.uk/11/4/4.html

Moges, A., & Holden, N. M. (2008). Soil fertility in relation to slope position and agricultural land use: a case study of Umbulo Catchment in southern Ethiopia. *Environmental management*, *42*(5), 753–763. doi:10.1007/s00267-008-9157-8

Myers Madeira, E. C. (2008). Policies to Reduce Emissions from Deforestation and Degradation (redd) in Developing Countries Policies to Reduce Emissions from Deforestation and Degradation (redd) in Developing Countries. *Change*, (December 2007), 84. Retrieved from http://unfccc.int/resource/docs/2008/smsn/ngo/016.pdf

Nava-Tablada, M. E., & Martínez-Camarillo, E. (2012). International migration and change in land use in Bella Esperanza, Veracruz. *Tropical and Subtropical Agroecosystems*, *15*, S21–S29. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84871062096&partnerID=40&md5=333ee4258fda4dd29ecbf054ac519b07

Nelson, G., De Pinto, A., Harris, V., & Stone, S. (2004). Land Use and Road Improvements: A Spatial Perspective. *International Regional Science Review*, *27*(3), 297–325. doi:10.1177/0160017604266028

Oxfam, M. (2002). Poverty in Your Coffee Cup. *Boston: Oxfam America*. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Poverty+in+your+coffee+cup#4

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P. (2003). Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. *Annals of the Association of American Geographers*, *93*(2), 314–337. doi:10.1111/1467-8306.9302004

Peierls, B. L., Caraco, N. F., Pace, M. L., & Cole, J. J. (1991). Human influence on river nitrogen. *Nature*, 386–387.

Perfecto, I., Vandermeer, J., Mas, A., & Pinto, L. S. (2005). Biodiversity, yield, and shade coffee certification. *Ecological Economics*, *54*(4), 435–446. doi:10.1016/j.ecolecon.2004.10.009

Provencher, B., & Baerenklau, K. A. (2005). An exposition of structural estimation of discrete dynamic decision processes. In *Applications of Simulation Methods in Environmental and Resource Economics.* (pp. 295–315). Netherlands: Springer.

Puri, J. (2006). Factors affecting agricultural expansion in forest reserves of Thailand: The role of population and roads. Retrieved from http://drum.lib.umd.edu/bitstream/1903/3481/1/umi-umd-3308.pdf

Redman, C. L., Grove, J. M., & Kuby, L. H. (2004). Integrating Social Science into the Long-Term Ecological Research (LTER) Network: Social Dimensions of Ecological Change and Ecological Dimensions of Social Change. *Ecosystems*, *7*, 161–171. doi:10.1007/s10021-003-0215-z

Robertson, R. D., Nelson, G. C., & De Pinto, A. (2009). Investigating the predictive capabilities of discrete choice models in the presence of spatial effects. *Papers in Regional Science*, *88*(2), 367–388. doi:10.1111/j.1435-5957.2009.00247.x

Romero, K., Houston, J. E., & Epperson, J. E. (2006). Diversification in Low-Grade Coffee-Growing Areas of Veracruz, Mexico: Market Possibilities. *Journal of Food Distribution Research*, *37*(1), 143–148.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines : An Empirical Model of Harold Zurcher. *Econometrica: Journal of the Econometric Society*, *55*(5), 999–1033. Retrieved from http://www.jstor.org/stable/10.2307/1911259

Rust, J. (1994). Structural Estimation of Markov Decision Processes. In *Handbook of Econometrics* (Vol. 4, pp. 3081–3143). Elsevier. doi:10.1016/S1573-4412(05)80020-0

SAGARPA. (2006). *Servicio de Información y Estadística Agroalimentaria y Pesquera*. Mexico City. Retrieved from http://www.siap.gob.mx/index.php?option=com_content&view=article&id=44&Itemid=378

SAGARPA. (2012). *Sistema de Información Agropecuaria de Consulta in Servicio de Información y Estadística Agroalimentaria y Pesquera.* Mexico City: Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación. Retrieved from http://www.siap.gob.mx/index.php?option=com_content&view=article&id=181&Itemid=426

Schnier, K. E., & Felthoven, R. G. (2011). Accounting for Spatial Heterogeneity and Autocorrelation in Spatial Discrete Choice Models: Implications for Behavioral Predictions. *Land Economics*, *87*(3), 382–402. Retrieved from http://le.uwpress.org/cgi/content/abstract/87/3/382

Schroth, G. (2004). *Agroforestry and biodiversity conservation in tropical landscapes*. *Island Press* (p. 540). doi:10.1007/s10457-006-9011-z

SEMARNAP. (1998). Mapa de suelos dominantes de la República Mexicana. Secretaría del Medio Ambiente, Recursos Naturales y Pesca.

SEMARNAT. (2005). *Informe de la situación del medio ambiente en México; compendio de estadísticas ambientales.* Mexico City.

Shanker, C., & Solanki, K. R. (2000). Agroforestry: an ecofriendly land-use system for insect management. *Outlook on Agriculture*. doi:10.5367/000000000101293095

Stryhn, H., & Christensen, J. (2003). Confidence intervals by the profile likelihood method, with applications in veterinary epidemiology. In *Proceedings of the 10th International Symposium on Veterinary Epidemiology and Economics* (Vol. 0, p. 208). Viña del Mar, Chile. Retrieved from http://people.upei.ca/hstryhn/stryhn208.pdf

Susaki, J., & Shibasaki, R. (2000). Maximum Likelihood Method Modified in Estimating a Prior Probability in Improving Misclassication Errors. *International Archives of Photogrammetry and Remote Sensing.* Amsterdam.

Swallow, B., Boffa, J., & Scherr, S. J. (2006). The potential for agroforestry to contribute to the conservation and enhancement of landscape biodiversity. In D. Garrity, A. Okono, M. Grayson, & S. Parrott (Eds.), *World Agroforestry into the Future* (pp. 95–101). World Agroforesty Centre.

Tauchen, G., & Hussey, R. (1991). Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models. *Econometrica*, *59*(2), 371–396. doi:10.2307/2938261

Tergas, L. E., & Sanchez, P. A. (1979). *Produccion de pastos en suelos acidos de los tropicos.* (Serie 03SG.).

Tian, G., Salako, F. K., Kolawole, G. O., & Kang, B. T. (1999). An Improved Cover Crop-Fallow System For Sustainable Management Of Low Activity Clay Soils Of The Tropics. *Soil Science*. doi:10.1097/00010694-199909000-00007

Train, K. (2009). *Discrete choice methods with simulation*. (C. U. Press, Ed.)*New York* (2nd. ed., Vol. 47). New York: Cambridge University Press. doi:10.1016/S0898-1221(04)90100-9

USDA. (2008). Soil Quality Indicators. Natural Resource Conservation Service. Retrieved August 17, 2010, from http://soils.usda.gov/sqi/assessment/files/infiltration_sq_physical_indicator_sheet.pdf

Van de Koppel, J., Rietkerk, M., & Weissing, F. J. (1997). Catastrophic vegetation shifts and soil degradation in terrestrial grazing systems. *Trends in ecology & evolution (Personal edition)*, *12*(9), 352–356. doi:10.1016/S0169-5347(97)01133-6

Verburg, P. H., Veldkamp, T. A., & Bouma, J. (1999). Land use change under conditions of high population pressure: the case of Java. *Global Environmental Change*, *9*(4), 303–312. doi:10.1016/S0959-3780(99)00175-2

Wei, X., Fu, X., Horton, R., & Shao, M. (2010). Changes in soil organic carbon and total nitrogen after 28 years grassland afforestation: effects of tree species, slope position, and soil order. *Plant and Soil*. doi:10.1007/s11104-009-0243-3

Wilkinson, G. G. (2005). Results and Implications of a Study of Fifteen Years of Satellite Image Classification Experiments. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(3), 433–440. doi:10.1109/TGRS.2004.837325

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., … Steinberg, D. (2007). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37. doi:10.1007/s10115-007-0114-2

Yang, H., Peng, J., Xia, B., & Zhang, D. (2013). An improved EM algorithm for remote sensing classification. *Chinese Science Bulletin*, *58*(9), 1060–1071. doi:10.1007/s11434-012-5485-4

Young, R., & Johnson, D. (2010). "Imputing the Missing Y"s: Implications for Survey Producers and Survey Users. *64th Annual Conference of the American …*, 6242–6248. Retrieved from http://www.amstat.org/Sections/Srms/Proceedings/y2010/Files/400142.pdf

Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2002). strucchange : An R Package for Testing for Structural Change in Linear Regression Models. *Journal Of Statistical Software*, *7*(2), 1–38. Retrieved from http://www.jstatsoft.org/v07/i02/

Zhai, D. (2007). *A note on the expectation-maximization (EM) algorithm.* Retrieved from http://times.cs.uiuc.edu/course/410s13/em-note.pdf

Zia, A. (2012). Land Use Adaptation to Climate Change: Economic Damages from Land-Falling Hurricanes in the Atlantic and Gulf States of the USA, 1900–2005. *Sustainability*, *4*, 917–932. doi:10.3390/su4050917